secondtake

December 13, 2024

1 Project: California Equity Research

1.0.1 Data: postgis db `calif_equity` with california climate investment and california enviroscreen data

Goal: Analyze the relationship between climate investment and environmental justice in California

This notebook: second take

Author: dpadams

Date: 2024-11-24

[2]:
```
# Importing the necessary libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import os
```

[3]:
```
## set directory
import os
os.chdir('/home/dadams/Repos/california_equity_git')
```

[4]:
```
# read in the data
data = pd.read_csv('data_raw/cci_programs_data.csv', low_memory=False)
```

[5]:
```
data.columns
```

[5]:
```
Index(['Project IDNumber', 'Reporting Cycle Name', 'Agency Name',
       'Program Name', 'Program Description', 'Sub Program Name',
       'Record Type', 'Project Name', 'Project Type', 'Project Description',
       …
       'Net Density  DUA', 'Applicants  Assisted', 'Invasive Cover 12 Months',
       'Invasive Cover 36 Months', 'Project Acreage', 'IS IAE',
       'Intermediary Admin Expenses Calc', 'PRIMARY_FUNDING_RECIPIENT_TYPE',
       'TRIBAL AFFILIATION', 'PROJECT PARTNERS'],
      dtype='object', length=127)
```

[6]:
```
import geopandas as gpd

# Load the shapefile
shapefile_path = '/home/dadams/Repos/california_equity_git/
 ↪california_enviroscreen/calif_enviroscreen_shape/CES4 Final Shapefile.shp'
gdf = gpd.read_file(shapefile_path)

# Print the head of the GeoDataFrame
print(gdf.head())
```

```
         Tract    ZIP         County    ApproxLoc  TotPop19    CIscore  \
0  6.083002e+09  93454  Santa Barbara  Santa Maria      4495  36.019653
1  6.083002e+09  93455  Santa Barbara  Santa Maria     13173  37.030667
2  6.083002e+09  93454  Santa Barbara  Santa Maria      2398  31.213140
3  6.083002e+09  93455  Santa Barbara       Orcutt      4496   6.639331
4  6.083002e+09  93455  Santa Barbara       Orcutt      4008  14.022852

    CIscoreP     Ozone     OzoneP     PM2_5  …  Elderly65  Hispanic  \
0  69.162885  0.034190  10.566273  7.567724  …    12.5028   68.9210
1  70.637922  0.035217  11.561917  7.624775  …     5.3519   78.6229
2  61.069087  0.034190  10.566273  7.548835  …    12.8857   65.7214
3   5.988401  0.036244  13.615432  7.660570  …    14.4128   22.9537
4  23.121533  0.036244  13.615432  7.663210  …    18.8872   33.4082

     White  AfricanAm  NativeAm  OtherMult    Shape_Leng    Shape_Area  \
0  20.8899     0.4004    0.2670     1.3126   6999.357689  2.847611e+06
1  13.2240     2.5051    0.0000     0.9489  19100.578232  1.635292e+07
2  30.6088     0.9591    0.0000     2.1685   4970.985897  1.352329e+06
3  69.1948     0.9342    0.7117     2.5356   6558.956012  2.417717e+06
4  59.7804     0.6986    1.4721     1.3723   6570.368730  2.608422e+06

     AAPI                                           geometry
0  8.2091  POLYGON ((-39795.07 -341919.191, -38126.384 -3…
1  4.6990  POLYGON ((-39795.07 -341919.191, -39803.632 -3…
2  0.5421  POLYGON ((-38115.747 -341130.248, -38126.384 -…
3  3.6699  POLYGON ((-37341.662 -348530.437, -37252.307 -…
4  3.2685  POLYGON ((-39465.107 -348499.262, -38244.305 -…

[5 rows x 67 columns]
```

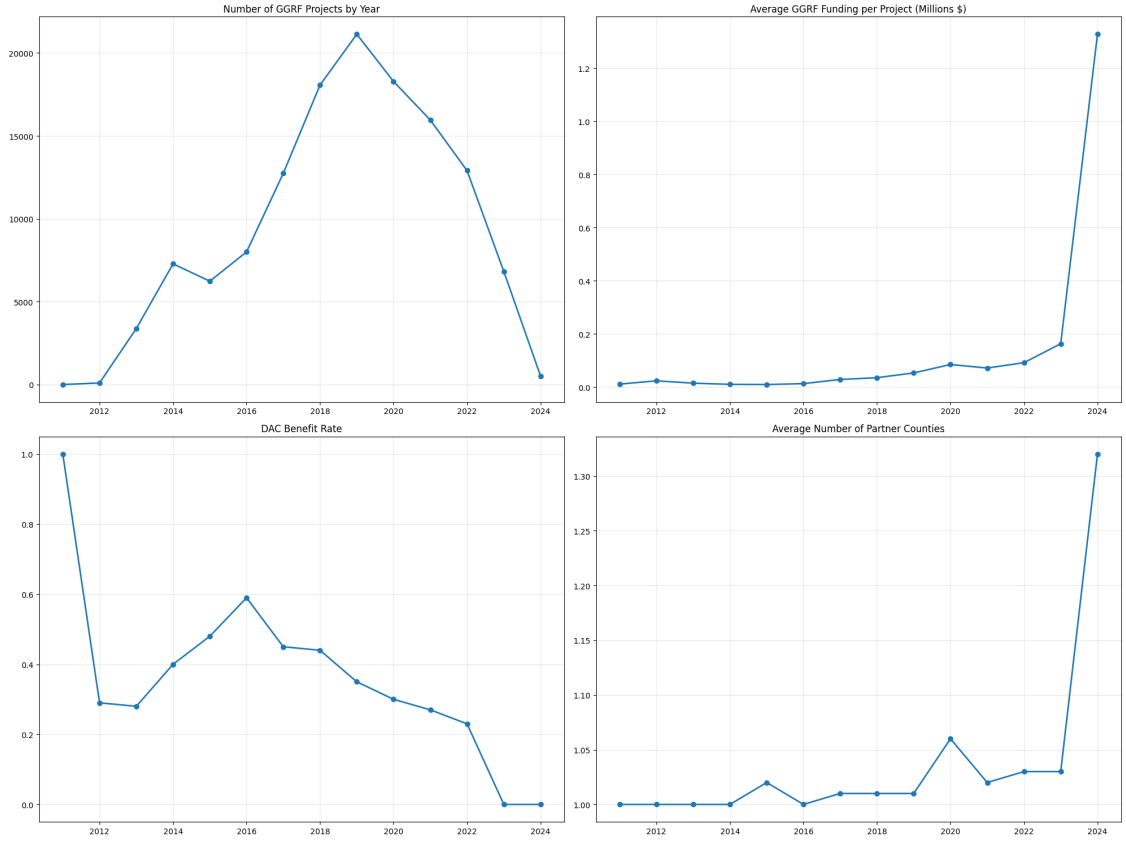
[7]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# Basic cleaning
data['Date Operational'] = pd.to_datetime(data['Date Operational'])
data = data[
```

```
    (data['Date Operational'] >= '2010-01-01') &
    (data['Date Operational'] <= '2024-11-01')
].copy()

# Remove rows with no GGRF funding
data = data.dropna(subset=['Total Program GGRFFunding'])

# Add derived columns
data['Year'] = data['Date Operational'].dt.year
data['is_multi_county'] = data['County'].str.contains(',', na=False)
data['partnership_size'] = data['County'].str.count(',').fillna(0) + 1

# Quick validation
print(f"Total GGRF Funding: ${data['Total Program GGRFFunding'].sum()/1e9:.
 ↪2f}B")
print(f"Number of projects: {len(data)}")
```

```
Total GGRF Funding: $8.13B
Number of projects: 131428
```

[8]:
```
# Temporal analysis of GGRF funding
temporal = data.groupby('Year').agg({
   'Total Program GGRFFunding': ['count', 'sum', 'mean'],
   'Is Benefit Disadvantaged Communities': 'mean',
   'is_multi_county': ['count', 'mean'],
   'partnership_size': 'mean'
}).round(2)

# Visualize key metrics
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(20, 15))

# Projects per year
ax1.plot(temporal.index, temporal[('Total Program GGRFFunding', 'count')],
        marker='o', linewidth=2)
ax1.set_title('Number of GGRF Projects by Year')
ax1.grid(True, alpha=0.3)

# Average funding per project
ax2.plot(temporal.index, temporal[('Total Program GGRFFunding', 'mean')]/1e6,
        marker='o', linewidth=2)
ax2.set_title('Average GGRF Funding per Project (Millions $)')
ax2.grid(True, alpha=0.3)

# DAC benefit rate
ax3.plot(temporal.index, temporal[('Is Benefit Disadvantaged Communities',␣
 ↪'mean')],
        marker='o', linewidth=2)
```

```
ax3.set_title('DAC Benefit Rate')
ax3.grid(True, alpha=0.3)

# Multi-county projects
ax4.plot(temporal.index, temporal[('partnership_size', 'mean')],
        marker='o', linewidth=2)
ax4.set_title('Average Number of Partner Counties')
ax4.grid(True, alpha=0.3)

plt.tight_layout()
plt.show()

print("Key metrics by period:")
print(temporal)
```




```
Key metrics by period:
     Total Program GGRFFunding                          \
                         count         sum         mean
Year
2011                         1       11500     11500.00
2012                        98     2328417     23759.36
```

2013	3379	50726172	15012.18
2014	7281	76042854	10444.01
2015	6235	61034095	9788.95
2016	8001	105685277	13209.01
2017	12745	368260901	28894.54
2018	18071	641426028	35494.77
2019	21131	1131373505	53540.94
2020	18281	1555581678	85092.81
2021	15957	1147697486	71924.39
2022	12906	1187728536	92029.18
2023	6825	1114799921	163340.65
2024	517	686268968	1327406.13

	Is Benefit Disadvantaged Communities		is_multi_county	\
	mean		count	mean
Year				
2011	1.00		1	0.00
2012	0.29		98	0.00
2013	0.28		3379	0.00
2014	0.40		7281	0.00
2015	0.48		6235	0.01
2016	0.59		8001	0.00
2017	0.45		12745	0.00
2018	0.44		18071	0.00
2019	0.35		21131	0.00
2020	0.30		18281	0.01
2021	0.27		15957	0.01
2022	0.23		12906	0.01
2023	0.00		6825	0.01
2024	0.00		517	0.07

	partnership_size
	mean
Year	
2011	1.00
2012	1.00
2013	1.00
2014	1.00
2015	1.02
2016	1.00
2017	1.01
2018	1.01
2019	1.01
2020	1.06
2021	1.02
2022	1.03
2023	1.03
2024	1.32

[9]:
```
print("2024 Projects by Program:")
print(data[data['Year'] == 2024].groupby('Program Name').agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean']
}).round(2))

print("\nLargest 2024 Projects:")
print(data[data['Year'] == 2024].nlargest(5, 'Total Program GGRFFunding')[
    ['Program Name', 'County', 'Total Program GGRFFunding', 'Date Operational']
])
```

```
2024 Projects by Program:
                                                    Total Program GGRFFunding  \
                                                                        count
Program Name
Affordable Housing and Sustainable Communities …                            8
Climate Adaptation and Resiliency Program                                   1
Community Air Protection                                                  220
Fire Prevention Program                                                    12
Fluorinated Gases Emission Reduction Incentives                            15
Food Production Investment Program                                         30
Forest Carbon Plan Implementation                                           5
Forest Health Program                                                      20
Low Carbon Transit Operations Program                                      18
Low Carbon Transportation                                                   4
Safe and Affordable Drinking Water Fund                                    10
Transformative Climate Communities                                         10
Transit and Intercity Rail Capital Program                                 29
Urban and Community Forestry Program                                      132
Waste Diversion                                                            1
Wetlands and Watershed Restoration                                         2


                                                        sum          mean
Program Name
Affordable Housing and Sustainable Communities …  176615877  22076984.62
Climate Adaptation and Resiliency Program            299000    299000.00
Community Air Protection                           80955408    367979.13
Fire Prevention Program                             7806649    650554.08
Fluorinated Gases Emission Reduction Incentives     1000001     66666.73
Food Production Investment Program                 70824290   2360809.67
Forest Carbon Plan Implementation                   1108131    221626.20
Forest Health Program                              57230331   2861516.55
Low Carbon Transit Operations Program              11165512    620306.22
Low Carbon Transportation                          10744732   2686183.00
Safe and Affordable Drinking Water Fund            10457866   1045786.60
Transformative Climate Communities                 38277301   3827730.10
Transit and Intercity Rail Capital Program        189696000   6541241.38
Urban and Community Forestry Program               18561997    140621.19
```

```
Waste Diversion                                        3950527   3950527.00
Wetlands and Watershed Restoration                     7575346   3787673.00

Largest 2024 Projects:
                                                 Program Name         County  \
117879         Transit and Intercity Rail Capital Program        Alameda
90922          Transit and Intercity Rail Capital Program    Los Angeles
136661  Affordable Housing and Sustainable Communities…    Los Angeles
141400  Affordable Housing and Sustainable Communities…  San Francisco
100763  Affordable Housing and Sustainable Communities…  San Francisco

        Total Program GGRFFunding Date Operational
117879                  107100000       2024-03-01
90922                    40000000       2024-09-23
136661                   29889806       2024-04-01
141400                   29269952       2024-08-01
100763                   25424799       2024-01-01
```

[10]:
```python
# Compare 2023 vs 2024 by program
years_comparison = data[data['Year'].isin([2023, 2024])].groupby(['Year',␣
 ↪'Program Name']).agg({
    'Total Program GGRFFunding': ['count', 'sum'],
    'Is Benefit Disadvantaged Communities': 'mean'
}).round(2)

print("2023 vs 2024 Program Comparison:")
print(years_comparison)

# Calculate percent changes in key metrics
print("\nPercent Changes 2023-2024:")
metrics_2023 = data[data['Year'] == 2023].agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Is Benefit Disadvantaged Communities': 'mean'
})

metrics_2024 = data[data['Year'] == 2024].agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Is Benefit Disadvantaged Communities': 'mean'
})

pct_change = ((metrics_2024 - metrics_2023) / metrics_2023 * 100).round(2)
print(pct_change)
```

```
2023 vs 2024 Program Comparison:
                                                      Total Program
GGRFFunding  \
count
Year Program Name
```

2023 Affordable Housing and Sustainable Communities …
20
Climate Adaptation and Resiliency Program
6
Climate Change Research Program
3
Climate Ready Program
4
Climate Resilience Planning
36
Climate Smart Agriculture
363
Community Air Protection
1926
Fire Prevention Program
11
Food Production Investment Program
20
Forest Carbon Plan Implementation
83
Forest Health Program
27
Funding Agricultural Replacement Measures for E…
539
Low Carbon Transit Operations Program
99
Low Carbon Transportation
3187
Low-Income Weatherization Program
26
SB 1383 Local Assistance Grant Program
12
Safe and Affordable Drinking Water Fund
22
Sustainable Agricultural Lands Conservation Pro…
15
Training and Workforce Development Program
136
Transformative Climate Communities
142
Transit and Intercity Rail Capital Program
14
Urban Greening Program
42
Waste Diversion
18
Water-Energy Efficiency
51

```
     Wetlands and Watershed Restoration
1
     Woodsmoke Reduction Program
22
2024 Affordable Housing and Sustainable Communities …
8
     Climate Adaptation and Resiliency Program
1
     Community Air Protection
220
     Fire Prevention Program
12
     Fluorinated Gases Emission Reduction Incentives
15
     Food Production Investment Program
30
     Forest Carbon Plan Implementation
5
     Forest Health Program
20
     Low Carbon Transit Operations Program
18
     Low Carbon Transportation
4
     Safe and Affordable Drinking Water Fund
10
     Transformative Climate Communities
10
     Transit and Intercity Rail Capital Program
29
     Urban and Community Forestry Program
132
     Waste Diversion
1
     Wetlands and Watershed Restoration
2
```

```
                                                                  \
                                                                sum
Year Program Name
2023 Affordable Housing and Sustainable Communities …  334283890
     Climate Adaptation and Resiliency Program            3374200
     Climate Change Research Program                       396119
     Climate Ready Program                                4368244
     Climate Resilience Planning                          5041046
     Climate Smart Agriculture                           43068601
     Community Air Protection                            45727103
     Fire Prevention Program                              9456431
```

```
      Food Production Investment Program                    42815128
      Forest Carbon Plan Implementation                     22524961
      Forest Health Program                                 37902075
      Funding Agricultural Replacement Measures for E…      40805691
      Low Carbon Transit Operations Program                114507056
      Low Carbon Transportation                            169989768
      Low-Income Weatherization Program                      4226240
      SB 1383 Local Assistance Grant Program                 5969666
      Safe and Affordable Drinking Water Fund               11621901
      Sustainable Agricultural Lands Conservation Pro…      24711311
      Training and Workforce Development Program             7332437
      Transformative Climate Communities                    40263894
      Transit and Intercity Rail Capital Program           118568000
      Urban Greening Program                                26722800
      Waste Diversion                                         213843
      Water-Energy Efficiency                                  73165
      Wetlands and Watershed Restoration                      743216
      Woodsmoke Reduction Program                              93135
2024  Affordable Housing and Sustainable Communities …     176615877
      Climate Adaptation and Resiliency Program               299000
      Community Air Protection                              80955408
      Fire Prevention Program                                7806649
      Fluorinated Gases Emission Reduction Incentives        1000001
      Food Production Investment Program                    70824290
      Forest Carbon Plan Implementation                      1108131
      Forest Health Program                                 57230331
      Low Carbon Transit Operations Program                 11165512
      Low Carbon Transportation                             10744732
      Safe and Affordable Drinking Water Fund               10457866
      Transformative Climate Communities                    38277301
      Transit and Intercity Rail Capital Program           189696000
      Urban and Community Forestry Program                  18561997
      Waste Diversion                                        3950527
      Wetlands and Watershed Restoration                     7575346

                                                        Is Benefit Disadvantaged
Communities
        mean
Year  Program Name
2023  Affordable Housing and Sustainable Communities …
0.00
      Climate Adaptation and Resiliency Program
0.00
      Climate Change Research Program
0.00
      Climate Ready Program
0.00
      Climate Resilience Planning
```

```
0.00
     Climate Smart Agriculture
0.00
     Community Air Protection
0.00
     Fire Prevention Program
0.00
     Food Production Investment Program
0.00
     Forest Carbon Plan Implementation
0.00
     Forest Health Program
0.00
     Funding Agricultural Replacement Measures for E…
0.00
     Low Carbon Transit Operations Program
0.00
     Low Carbon Transportation
0.00
     Low-Income Weatherization Program
0.00
     SB 1383 Local Assistance Grant Program
0.00
     Safe and Affordable Drinking Water Fund
0.00
     Sustainable Agricultural Lands Conservation Pro…
0.00
     Training and Workforce Development Program
0.00
     Transformative Climate Communities
0.00
     Transit and Intercity Rail Capital Program
0.36
     Urban Greening Program
0.00
     Waste Diversion
0.00
     Water-Energy Efficiency
0.06
     Wetlands and Watershed Restoration
0.00
     Woodsmoke Reduction Program
0.00
2024 Affordable Housing and Sustainable Communities …
0.00
     Climate Adaptation and Resiliency Program
0.00
     Community Air Protection
```

```
0.00
     Fire Prevention Program
0.00
     Fluorinated Gases Emission Reduction Incentives
0.00
     Food Production Investment Program
0.00
     Forest Carbon Plan Implementation
0.00
     Forest Health Program
0.00
     Low Carbon Transit Operations Program
0.00
     Low Carbon Transportation
0.00
     Safe and Affordable Drinking Water Fund
0.00
     Transformative Climate Communities
0.00
     Transit and Intercity Rail Capital Program
0.00
     Urban and Community Forestry Program
0.00
     Waste Diversion
0.00
     Wetlands and Watershed Restoration
0.00

Percent Changes 2023-2024:
       Total Program GGRFFunding  Is Benefit Disadvantaged Communities
count                     -92.42                                   NaN
sum                       -38.44                                   NaN
mean                      712.66                                -100.0
```

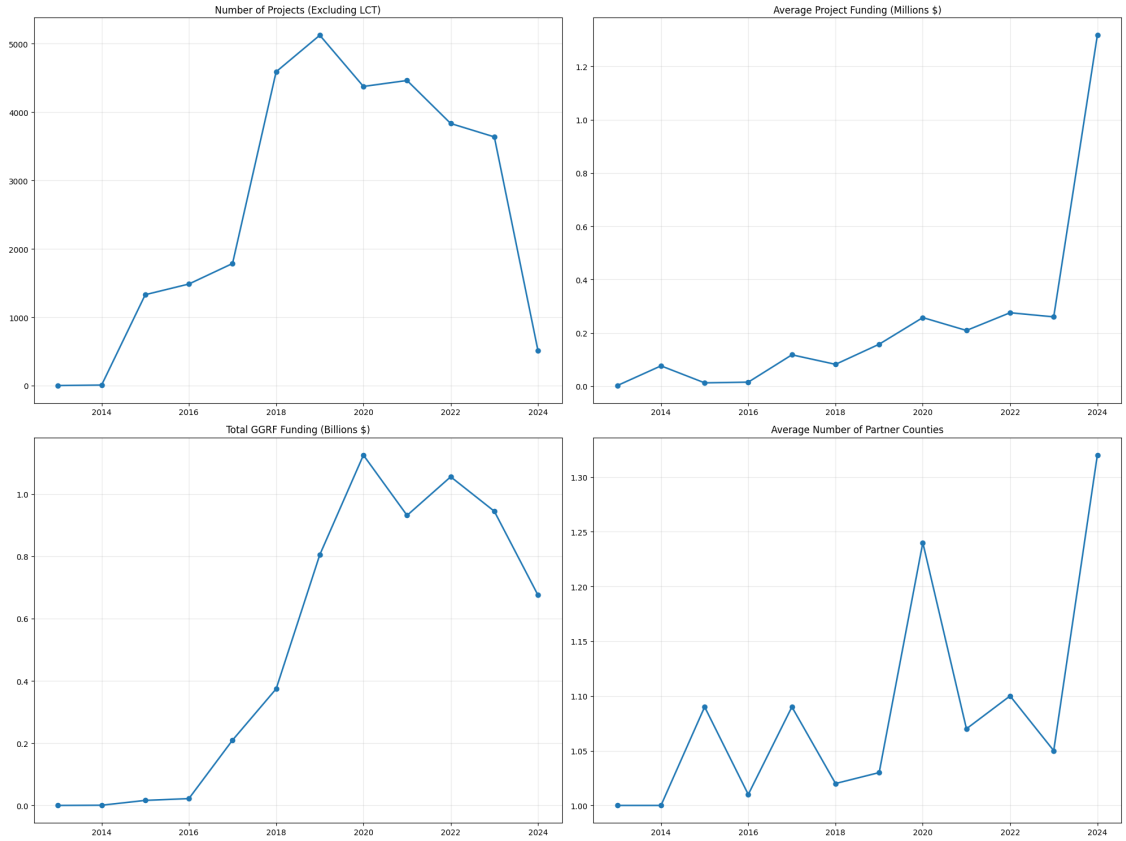
[11]:
```python
# Filter out Low Carbon Transportation
data_filtered = data[data['Program Name'] != 'Low Carbon Transportation'].copy()

# Recalculate temporal analysis
temporal_filtered = data_filtered.groupby('Year').agg({
   'Total Program GGRFFunding': ['count', 'sum', 'mean'],
   'Is Benefit Disadvantaged Communities': 'mean',
   'is_multi_county': ['count', 'mean'],
   'partnership_size': 'mean'
}).round(2)

# Compare 2023-2024 without LCT
```

```
years_comparison = data_filtered[data_filtered['Year'].isin([2023, 2024])].
 ↪groupby(['Year', 'Program Name']).agg({
   'Total Program GGRFFunding': ['count', 'sum'],
   'Is Benefit Disadvantaged Communities': 'mean'
}).round(2)

print("2023-2024 Changes (excluding Low Carbon Transportation):")
changes = pd.DataFrame({
   '2023': temporal_filtered.loc[2023],
   '2024': temporal_filtered.loc[2024]
})
print(changes)
```

```
2023-2024 Changes (excluding Low Carbon Transportation):
                                                    2023          2024
Total Program GGRFFunding            count  3.638000e+03  5.130000e+02
                                     sum    9.448102e+08  6.755242e+08
                                     mean   2.597059e+05  1.316811e+06
Is Benefit Disadvantaged Communities mean   0.000000e+00  0.000000e+00
is_multi_county                      count  3.638000e+03  5.130000e+02
                                     mean   2.000000e-02  7.000000e-02
partnership_size                     mean   1.050000e+00  1.320000e+00
```

[12]:
```
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(20, 15))

# Projects count
ax1.plot(temporal_filtered.index,
         temporal_filtered[('Total Program GGRFFunding', 'count')],
         marker='o', linewidth=2)
ax1.set_title('Number of Projects (Excluding LCT)')
ax1.grid(True, alpha=0.3)

# Average funding
ax2.plot(temporal_filtered.index,
         temporal_filtered[('Total Program GGRFFunding', 'mean')]/1e6,
         marker='o', linewidth=2)
ax2.set_title('Average Project Funding (Millions $)')
ax2.grid(True, alpha=0.3)

# Total funding
ax3.plot(temporal_filtered.index,
         temporal_filtered[('Total Program GGRFFunding', 'sum')]/1e9,
         marker='o', linewidth=2)
ax3.set_title('Total GGRF Funding (Billions $)')
ax3.grid(True, alpha=0.3)

# Partnership size
```

```
ax4.plot(temporal_filtered.index,
         temporal_filtered[('partnership_size', 'mean')],
         marker='o', linewidth=2)
ax4.set_title('Average Number of Partner Counties')
ax4.grid(True, alpha=0.3)

plt.tight_layout()
plt.show()
```



1. **Project Volume Evolution**

- Growth phase: 2014-2019 (from ~0 to 5000+ projects)
- Plateau: 2019-2022 (~4000-5000 projects)
- Sharp decline: 2023-2024 (down to ~500 projects)

2. **Average Project Size**

- Relatively stable 2014-2023 ($0.1-0.3M per project)
- Dramatic increase in 2024 (to ~$1.3M per project)
- Suggests shift to fewer but larger projects

3. **Total GGRF Funding**

- Steady increase: 2014-2020 (reaching ~$1.1B)
- Recent decline: 2020-2024 (down to ~$0.67B)
- More stable pattern than project counts

4. **Partnership Trends**

- Generally stable at 1.0-1.1 partners until 2020
- Spike in 2020 (~1.24 partners)
- New peak in 2024 (~1.32 partners)
- Suggests increasing regional collaboration

[13]:
```
# 1. Program Scale Analysis
program_scale = data_filtered.groupby('Program Name').agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Is Benefit Disadvantaged Communities': 'mean',
    'Total Project GHGReductions': 'sum'
}).round(2)

# Flatten column names
program_scale.columns = ['project_count', 'total_funding', 'avg_funding',
                         'dac_rate', 'total_ghg']

# Calculate GHG efficiency
program_scale['ghg_per_dollar'] = program_scale['total_ghg'] /␣
 ↪program_scale['total_funding']

# Categorize programs by size
def categorize_program_size(mean_funding):
    if mean_funding > 10e6:  # 10M
        return 'Mega'
    elif mean_funding > 1e6:  # 1M
        return 'Large'
    elif mean_funding > 500e3:  # 500K
        return 'Medium'
    else:
        return 'Small'

program_scale['size_category'] = program_scale['avg_funding'].
 ↪apply(categorize_program_size)

print("Program Scale Distribution:")
print(program_scale['size_category'].value_counts())

# 2. Geographic Analysis
geographic_dist = data_filtered.groupby('County').agg({
    'Total Program GGRFFunding': ['count', 'sum'],
    'Is Benefit Disadvantaged Communities': 'mean',
    'Total Project GHGReductions': 'sum'
```

```
})

geographic_dist.columns = ['project_count', 'total_funding', 'dac_rate',␣
 ↪'total_ghg']

# Calculate concentration metrics
total_funding = geographic_dist['total_funding'].sum()
top_5_counties = geographic_dist['total_funding'].nlargest(5)
concentration = (top_5_counties.sum() / total_funding) * 100

print("\nGeographic Concentration:")
print(f"Top 5 counties account for {concentration:.1f}% of funding")

# 3. Print key findings
print("\nProgram Categories by Total Funding (Billions $):")
size_summary = program_scale.groupby('size_category')['total_funding'].sum().
 ↪sort_values(ascending=False)/1e9
print(size_summary.round(2))

# Show largest programs
print("\nLargest Programs (by total funding):")
print(program_scale.nlargest(5, 'total_funding')[
    ['total_funding', 'project_count', 'avg_funding', 'dac_rate',␣
 ↪'ghg_per_dollar']
].round(2))
```

```
Program Scale Distribution:
size_category
Small     21
Large     11
Medium     5
Mega       1
Name: count, dtype: int64

Geographic Concentration:
Top 5 counties account for 39.3% of funding

Program Categories by Total Funding (Billions $):
size_category
Large     2.19
Small     2.00
Mega      1.19
Medium    0.78
Name: total_funding, dtype: float64

Largest Programs (by total funding):
                                                    total_funding  \
```

```
Program Name
Affordable Housing and Sustainable Communities …     1192203125
Low Carbon Transit Operations Program                   775906434
Transit and Intercity Rail Capital Program              771556000
Fire Prevention Program                                 596274123
Community Air Protection                                529523228

                                                    project_count  \
Program Name
Affordable Housing and Sustainable Communities …             93
Low Carbon Transit Operations Program                         766
Transit and Intercity Rail Capital Program                    135
Fire Prevention Program                                       600
Community Air Protection                                     5187

                                                    avg_funding  dac_rate  \
Program Name
Affordable Housing and Sustainable Communities …  12819388.44      0.16
Low Carbon Transit Operations Program               1012932.68      0.09
Transit and Intercity Rail Capital Program          5715229.63      0.19
Fire Prevention Program                              993790.20      0.00
Community Air Protection                             102086.61      0.00

                                                    ghg_per_dollar
Program Name
Affordable Housing and Sustainable Communities …              0.00
Low Carbon Transit Operations Program                         0.01
Transit and Intercity Rail Capital Program                    0.01
Fire Prevention Program                                       0.00
Community Air Protection                                      0.00
```

1. **Program Scale Distribution**

- Most programs (21) are "Small" scale
- 11 "Large" programs
- Only 1 "Mega" program (Affordable Housing at $1.19B)
- More balanced distribution than when including transportation subsidies

1. **Funding Allocation**

- Large programs: $2.19B total
- Small programs: $2.00B total
- Mega programs: $1.19B total
- Medium programs: $0.78B total
- Total GGRF funding: ~$6.16B

1. **Top Programs by Funding**

- Affordable Housing: $1.19B (93 projects)
- Low Carbon Transit: $776M (766 projects)
- Transit/Rail Capital: $772M (135 projects)

- Fire Prevention: $596M (600 projects)
- Community Air Protection: $530M (5,187 projects)

1. **Program Characteristics**

- Wide range in project counts (93 to 5,187)
- Average project sizes vary significantly:
  - Affordable Housing: $12.8M/project
  - Transit/Rail: $5.7M/project
  - Community Air Protection: $102K/project

1. **Geographic Distribution**

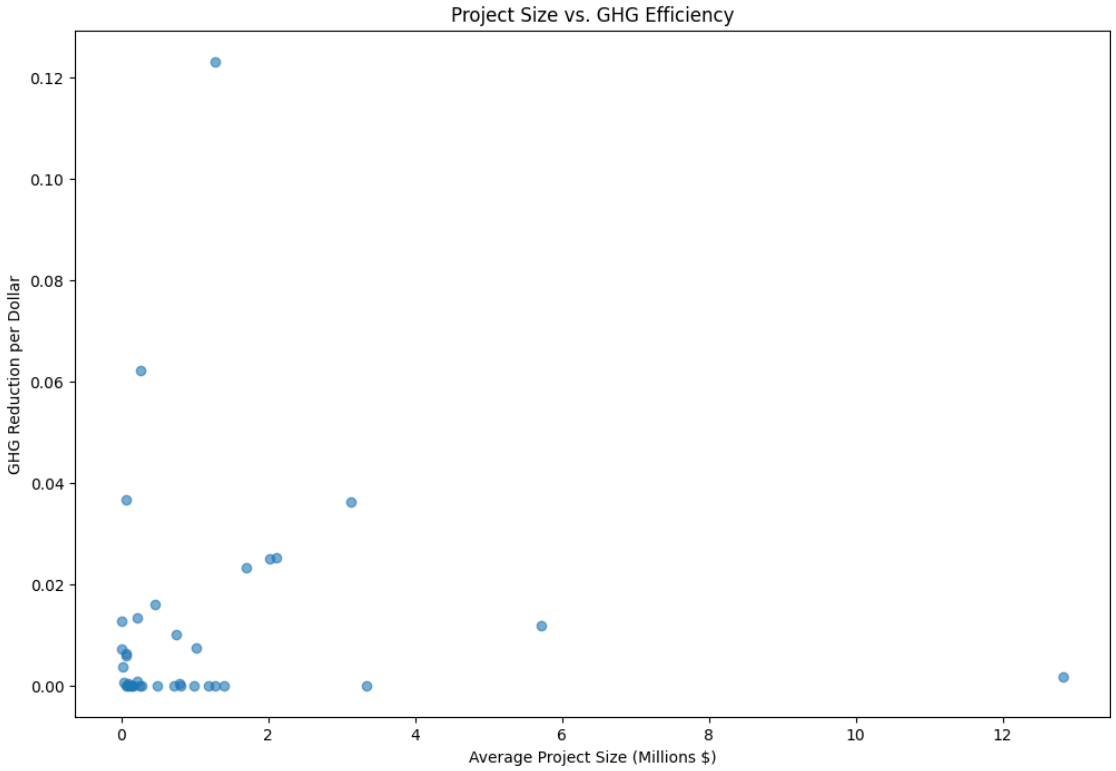
- Less concentrated than before
- Top 5 counties: 39.3% of funding (vs. previous 75.6%)
- Suggests more equitable geographic distribution

[14]:
```
# 1. DAC Benefits by Program Size
dac_by_size = program_scale.groupby('size_category').agg({
    'dac_rate': ['mean', 'min', 'max'],
    'total_funding': 'sum',
    'project_count': 'sum'
}).round(3)

# 2. Geographic Analysis
# Add region classification
def classify_region(county):
    if isinstance(county, str):  # Handle multi-county cases
        counties = county.split(',')
        county = counties[0].strip()

    urban_counties = ['Los Angeles', 'San Francisco', 'Alameda', 'San Diego',␣
 ↪'Orange']
    central_valley = ['Fresno', 'Kern', 'Kings', 'Madera', 'Merced', 'San␣
 ↪Joaquin', 'Stanislaus', 'Tulare']

    if county in urban_counties:
        return 'Urban'
    elif county in central_valley:
        return 'Central Valley'
    else:
        return 'Other'

geographic_dist['region'] = geographic_dist.index.map(classify_region)
regional_metrics = geographic_dist.groupby('region').agg({
    'total_funding': 'sum',
    'project_count': 'sum',
    'dac_rate': 'mean',
    'total_ghg': 'sum'
```

```
}).round(3)

# 3. Project Size vs GHG Efficiency
# Create scatter plot
plt.figure(figsize=(12, 8))
plt.scatter(program_scale['avg_funding']/1e6,
           program_scale['ghg_per_dollar'],
           alpha=0.6)
plt.xlabel('Average Project Size (Millions $)')
plt.ylabel('GHG Reduction per Dollar')
plt.title('Project Size vs. GHG Efficiency')

# 4. Multi-county Analysis
multi_county_data = data_filtered[data_filtered['County'].str.contains(',',␣
 ↪na=False)]
single_county_data = data_filtered[~data_filtered['County'].str.contains(',',␣
 ↪na=False)]

multi_vs_single = pd.DataFrame({
   'Multi-County': {
       'project_count': len(multi_county_data),
       'total_funding': multi_county_data['Total Program GGRFFunding'].sum(),
       'avg_funding': multi_county_data['Total Program GGRFFunding'].mean(),
       'dac_rate': multi_county_data['Is Benefit Disadvantaged Communities'].
 ↪mean(),
       'ghg_per_dollar': (multi_county_data['Total Project GHGReductions'].
 ↪sum() /
                          multi_county_data['Total Program GGRFFunding'].sum())
   },
   'Single-County': {
       'project_count': len(single_county_data),
       'total_funding': single_county_data['Total Program GGRFFunding'].sum(),
       'avg_funding': single_county_data['Total Program GGRFFunding'].mean(),
       'dac_rate': single_county_data['Is Benefit Disadvantaged Communities'].
 ↪mean(),
       'ghg_per_dollar': (single_county_data['Total Project GHGReductions'].
 ↪sum() /
                          single_county_data['Total Program GGRFFunding'].sum())
   }
})

print("1. DAC Benefits by Program Size:")
print(dac_by_size)

print("\n2. Regional Distribution:")
print(regional_metrics)
```

```
print("\n4. Multi-County vs Single-County Projects:")
print(multi_vs_single.round(3))

plt.show()
```

```
1. DAC Benefits by Program Size:
              dac_rate                 total_funding project_count
                  mean   min   max               sum           sum
size_category
Large            0.087  0.00  0.67    2185763970          1314
Medium           0.036  0.00  0.18     779274523           831
Mega             0.160  0.16  0.16    1192203125            93
Small            0.084  0.00  0.92    2003043460         28913

2. Regional Distribution:
                total_funding  project_count  dac_rate  total_ghg
region
Central Valley     1387932197          14496     0.104   25320641
Other              2289435192           9195     0.101   24839824
Urban              2482917689           7460     0.161   15727243

4. Multi-County vs Single-County Projects:
                Multi-County  Single-County
project_count   7.520000e+02   3.039900e+04
total_funding   5.210746e+08   5.639210e+09
avg_funding     6.929184e+05   1.855064e+05
dac_rate        1.090000e-01   2.730000e-01
ghg_per_dollar  1.200000e-02   1.100000e-02
```



1. **Efficiency Sweet Spot**

- Highest GHG efficiency (0.12-0.13 reduction per dollar) occurs in projects around $2M
- Secondary peak (0.06) in smaller projects under $1M
- Suggests optimal scale for GHG reduction isn't necessarily larger projects

1. **Size-Efficiency Relationship**

- No clear linear relationship
- Most projects cluster in lower ranges (both size and efficiency)
- Largest projects ($12M+) show relatively low GHG efficiency
- Suggests diminishing returns as projects scale up

1. **Distribution Pattern**

- Dense cluster of projects under $2M with varying efficiency
- Sparse distribution in higher project sizes
- Few projects achieve both large scale and high efficiency

Let's identify: 1. What programs occupy that sweet spot around $2M with high efficiency? 2. What characteristics do those efficient projects share? 3. Are there geographic patterns in the more efficient projects?

[15]:
```
# Let's identify the most GHG-efficient projects and analyze their␣
↪characteristics
```

```
# Add efficiency metric to original program data
program_scale['efficiency_tier'] = pd.qcut(program_scale['ghg_per_dollar'],
                                           q=4,
                                           labels=['Low', 'Medium-Low',␣
 ↪'Medium-High', 'High'])

print("1. Most GHG-Efficient Programs:")
high_efficiency = program_scale[program_scale['efficiency_tier'] == 'High'].
 ↪sort_values('ghg_per_dollar', ascending=False)
print(high_efficiency[['total_funding', 'avg_funding', 'ghg_per_dollar',␣
 ↪'dac_rate']].round(3))

# Analyze characteristics of high-efficiency programs
print("\n2. Characteristics of High-Efficiency Programs:")
print("\nSize Distribution:")
print(high_efficiency['size_category'].value_counts())

# Geographic analysis of high-efficiency programs
print("\n3. Geographic Distribution of High-Efficiency Projects:")
# Filter original data for these programs
high_eff_projects = data_filtered[
   data_filtered['Program Name'].isin(high_efficiency.index)
]
geographic_dist = high_eff_projects.groupby('County').agg({
   'Total Program GGRFFunding': ['count', 'sum'],
   'Total Project GHGReductions': 'sum'
}).round(2)

# Calculate efficiency by region
geographic_dist['ghg_per_dollar'] = (
   geographic_dist[('Total Project GHGReductions', 'sum')] /
   geographic_dist[('Total Program GGRFFunding', 'sum')]
)

print(geographic_dist.nlargest(5, 'ghg_per_dollar'))

# Create visualization of efficiency patterns
plt.figure(figsize=(15, 8))

plt.subplot(1, 2, 1)
# Size vs Efficiency for high performers
plt.scatter(high_efficiency['avg_funding']/1e6,
          high_efficiency['ghg_per_dollar'],
          alpha=0.6,
          s=100)
plt.xlabel('Average Project Size (Millions $)')
plt.ylabel('GHG Reduction per Dollar')
```

```
plt.title('High-Efficiency Programs: Size vs Performance')

# Add program labels
for idx, row in high_efficiency.iterrows():
   plt.annotate(idx[:20] + '...' if len(idx) > 20 else idx,
               (row['avg_funding']/1e6, row['ghg_per_dollar']),
               xytext=(5, 5), textcoords='offset points')

plt.subplot(1, 2, 2)
# DAC Rate vs Efficiency
plt.scatter(high_efficiency['dac_rate'],
          high_efficiency['ghg_per_dollar'],
          alpha=0.6,
          s=100)
plt.xlabel('DAC Benefit Rate')
plt.ylabel('GHG Reduction per Dollar')
plt.title('High-Efficiency Programs: Equity vs Performance')

plt.tight_layout()
plt.show()
```

```
1. Most GHG-Efficient Programs:
                                                  total_funding  \
Program Name
Sustainable Agricultural Lands Conservation Pro…      122424176
Climate Smart Agriculture                             338161549
Fluorinated Gases Emission Reduction Incentives         1000001
Low-Carbon Fuels Production Program                    12500000
Food Production Investment Program                    117791478
Wetlands and Watershed Restoration                     16171301
Forest Health Program                                 207621340
Waste Diversion                                       118571417
Renewable Energy for Agriculture Program               9500000
Woodsmoke Reduction Program                            7895909

                                                  avg_funding  \
Program Name
Sustainable Agricultural Lands Conservation Pro…  1275251.83
Climate Smart Agriculture                           257941.68
Fluorinated Gases Emission Reduction Incentives      66666.73
Low-Carbon Fuels Production Program                3125000.00
Food Production Investment Program                 2103419.25
Wetlands and Watershed Restoration                 2021412.62
Forest Health Program                              1701814.26
Waste Diversion                                     463169.60
Renewable Energy for Agriculture Program            211111.11
Woodsmoke Reduction Program                           8233.48
```

Program Name	ghg_per_dollar	dac_rate
Sustainable Agricultural Lands Conservation Pro…	0.123	0.01
Climate Smart Agriculture	0.062	0.05
Fluorinated Gases Emission Reduction Incentives	0.037	0.00
Low-Carbon Fuels Production Program	0.036	0.00
Food Production Investment Program	0.025	0.00
Wetlands and Watershed Restoration	0.025	0.00
Forest Health Program	0.023	0.00
Waste Diversion	0.016	0.09
Renewable Energy for Agriculture Program	0.013	0.00
Woodsmoke Reduction Program	0.013	0.00

2. Characteristics of High-Efficiency Programs:

Size Distribution:

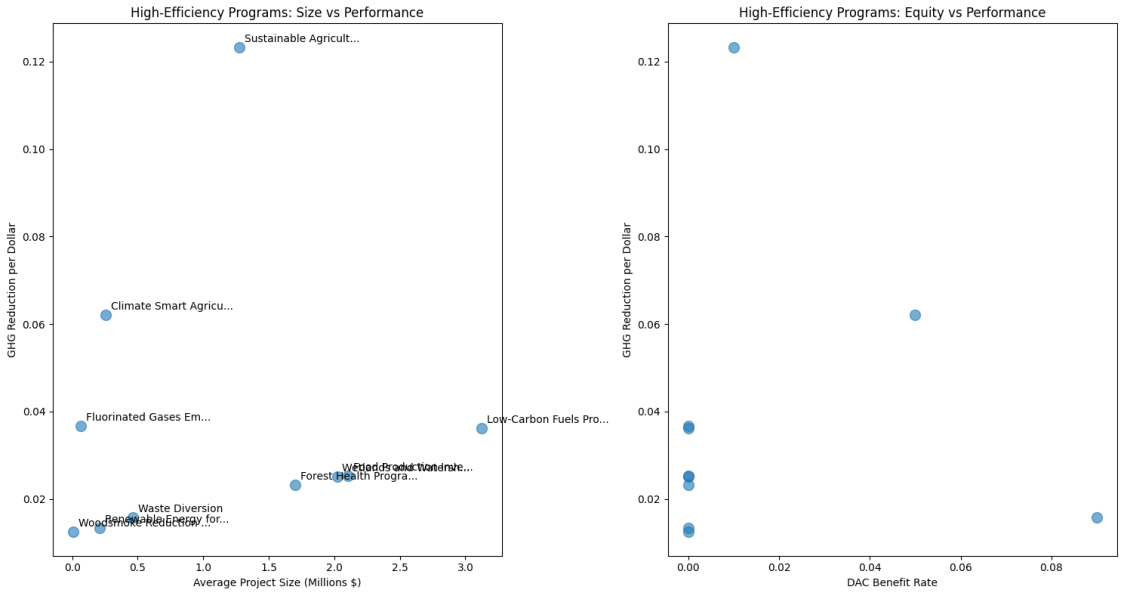
size_category	
Large	5
Small	5

Name: count, dtype: int64

3. Geographic Distribution of High-Efficiency Projects:

	Total Program GGRFFunding	
	count	sum
County		
Sierra	5	485667
Mono	10	2946879
Lassen	28	5510913
Calaveras	32	9678108
Los Angeles, Solano	1	212629

	Total Project GHGReductions	ghg_per_dollar
	sum	
County		
Sierra	730033	1.503155
Mono	1570341	0.532883
Lassen	2089977	0.379243
Calaveras	1857223	0.191899
Los Angeles, Solano	23680	0.111368

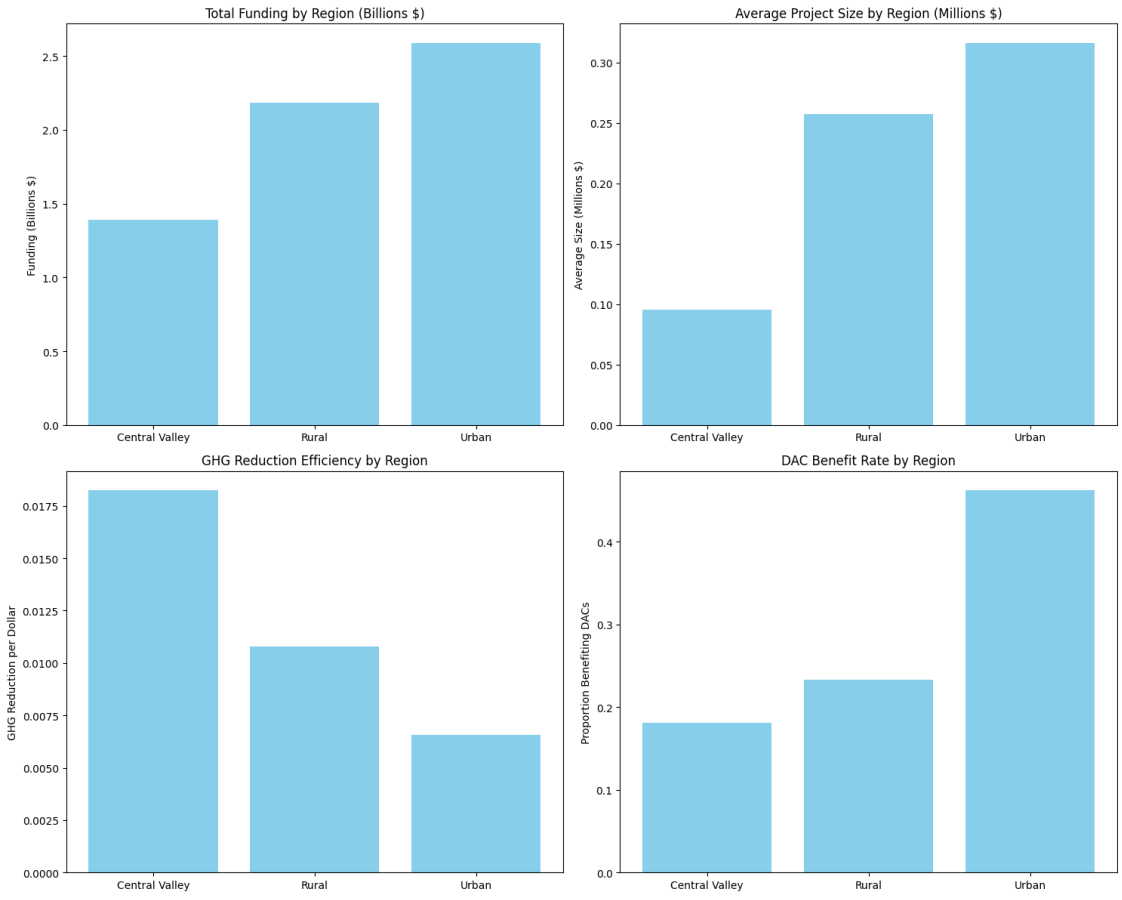


[16]:
```
# Define urban/rural classification
def classify_urban_rural(county):
    if isinstance(county, str):  # Handle multi-county cases
        counties = county.split(',')
        county = counties[0].strip()

    urban_counties = ['Los Angeles', 'San Francisco', 'Alameda', 'San Diego',␣
 ↪'Orange', 'Santa Clara']
    central_valley = ['Fresno', 'Kern', 'Kings', 'Madera', 'Merced', 'San␣
 ↪Joaquin', 'Stanislaus', 'Tulare']

    if county in urban_counties:
        return 'Urban'
    elif county in central_valley:
        return 'Central Valley'
    else:
        return 'Rural'

# Add classification to data
data_filtered['region_type'] = data_filtered['County'].map(classify_urban_rural)

# Analyze by region type
region_analysis = data_filtered.groupby('region_type').agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Total Project GHGReductions': ['sum', 'mean'],
    'Is Benefit Disadvantaged Communities': 'mean'
}).round(3)
```

```
# Calculate GHG efficiency by region
ghg_efficiency = (region_analysis[('Total Project GHGReductions', 'sum')] /
                  region_analysis[('Total Program GGRFFunding', 'sum')])

print("Urban/Rural Analysis Summary:")
summary_df = pd.DataFrame({
   'Project Count': region_analysis[('Total Program GGRFFunding', 'count')],
   'Total Funding (B)': region_analysis[('Total Program GGRFFunding', 'sum')]/
 ↪1e9,
   'Avg Project Size (M)': region_analysis[('Total Program GGRFFunding',␣
 ↪'mean')]/1e6,
   'GHG Efficiency': ghg_efficiency,
   'DAC Rate': region_analysis[('Is Benefit Disadvantaged Communities', 'mean')]
})
print(summary_df.round(3))

# Visualize key metrics
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(15, 12))

# 1. Total Funding by Region
ax1.bar(summary_df.index, summary_df['Total Funding (B)'], color='skyblue')
ax1.set_title('Total Funding by Region (Billions $)')
ax1.set_ylabel('Funding (Billions $)')

# 2. Average Project Size
ax2.bar(summary_df.index, summary_df['Avg Project Size (M)'], color='skyblue')
ax2.set_title('Average Project Size by Region (Millions $)')
ax2.set_ylabel('Average Size (Millions $)')

# 3. GHG Efficiency
ax3.bar(summary_df.index, summary_df['GHG Efficiency'], color='skyblue')
ax3.set_title('GHG Reduction Efficiency by Region')
ax3.set_ylabel('GHG Reduction per Dollar')

# 4. DAC Benefit Rate
ax4.bar(summary_df.index, summary_df['DAC Rate'], color='skyblue')
ax4.set_title('DAC Benefit Rate by Region')
ax4.set_ylabel('Proportion Benefiting DACs')

plt.tight_layout()
plt.show()

# Look at program distribution by region
print("\nTop Programs by Region (Funding in Millions $):")
program_by_region = pd.crosstab(
   data_filtered['Program Name'],
```

```
   data_filtered['region_type'],
   values=data_filtered['Total Program GGRFFunding'],
   aggfunc='sum'
)/1e6

print(program_by_region.sort_values('Urban', ascending=False).head(10).round(2))
```

```
Urban/Rural Analysis Summary:
                Project Count  Total Funding (B)  Avg Project Size (M)  \
region_type
Central Valley          14496              1.388                 0.096
Rural                    8474              2.183                 0.258
Urban                    8181              2.589                 0.316

                GHG Efficiency  DAC Rate
region_type
Central Valley           0.018     0.181
Rural                    0.011     0.233
Urban                    0.007     0.462
```



```
Top Programs by Region (Funding in Millions $):
region_type                                          Central Valley   Rural  \
Program Name
Transit and Intercity Rail Capital Program                    18.28   86.89
Affordable Housing and Sustainable Communities …             205.61  321.80
Low Carbon Transit Operations Program                         52.88  178.45
Community Air Protection                                     144.75  145.80
Transformative Climate Communities                            69.17   62.67
Urban Greening Program                                        18.91   58.50
Fire Prevention Program                                       51.09  493.55
Waste Diversion                                               22.71   50.54
Low-Income Weatherization Program                             62.87   57.02
Urban and Community Forestry Program                           7.55   13.80

region_type                                           Urban
Program Name
Transit and Intercity Rail Capital Program           666.38
Affordable Housing and Sustainable Communities …     664.79
Low Carbon Transit Operations Program                544.58
Community Air Protection                             238.97
Transformative Climate Communities                    70.89
Urban Greening Program                                61.36
Fire Prevention Program                               51.63
Waste Diversion                                       45.32
Low-Income Weatherization Program                     39.08
Urban and Community Forestry Program                  37.84
```

1. **GHG Efficiency Gap**

- Central Valley is ~3x more efficient than Urban areas (0.0175 vs 0.0065 GHG reduction per dollar)
- Rural areas also outperform Urban (0.010 vs 0.0065)
- Suggests climate dollars go further in non-urban areas

1. **Inverse DAC Pattern**

- Urban areas have highest DAC benefit rate (~0.45)
- Rural areas at ~0.23 DAC rate
- Central Valley lowest at ~0.18

1. **Investment Distribution**

- Urban areas get most funding ($2.5B)
- Rural second ($2.2B)
- Central Valley least ($1.4B)
- Average project sizes follow same pattern

[17]:
```
# Calculate efficiency ratios and examine program types
print("Program Types and Efficiency by Region:")
```

```
program_efficiency = data_filtered.groupby(['Program Name', 'region_type']).
 ↪agg({
    'Total Program GGRFFunding': 'sum',
    'Total Project GHGReductions': 'sum'
}).reset_index()

program_efficiency['efficiency'] = (program_efficiency['Total Project␣
 ↪GHGReductions'] /
                                    program_efficiency['Total Program␣
 ↪GGRFFunding'])

# Show top efficient programs by region
for region in ['Urban', 'Rural', 'Central Valley']:
    print(f"\n{region} Most Efficient Programs:")
    region_data = program_efficiency[program_efficiency['region_type'] ==␣
 ↪region]
    print(region_data.nlargest(5, 'efficiency')[
        ['Program Name', 'Total Program GGRFFunding', 'efficiency']
    ].round(4))
```

```
Program Types and Efficiency by Region:

Urban Most Efficient Programs:
                                        Program Name  \
100                       Woodsmoke Reduction Program
29   Fluorinated Gases Emission Reduction Incentives
61          Renewable Energy for Agriculture Program
89                                    Waste Diversion
32                Food Production Investment Program

     Total Program GGRFFunding  efficiency
100                        808      0.1027
29                      530239      0.0380
61                       42551      0.0290
89                    45316021      0.0203
32                    13221462      0.0129

Rural Most Efficient Programs:
                                           Program Name  \
69  Sustainable Agricultural Lands Conservation Pr…
49                Low-Carbon Fuels Production Program
28    Fluorinated Gases Emission Reduction Incentives
31                 Food Production Investment Program
37                              Forest Health Program

    Total Program GGRFFunding  efficiency
69                  108577417      0.1273
```

```
49                    5000000      0.0905
28                     469762      0.0352
31                   32710900      0.0298
37                  163034942      0.0285

Central Valley Most Efficient Programs:
                                          Program Name  \
68  Sustainable Agricultural Lands Conservation Pr…
14                           Climate Smart Agriculture
30                 Food Production Investment Program
87                                    Waste Diversion
59           Renewable Energy for Agriculture Program

    Total Program GGRFFunding  efficiency
68                    9744397      0.1253
14                  274120208      0.0751
30                   71859116      0.0255
87                   22712521      0.0252
59                    3761222      0.0180
```

This breakdown is really revealing about what works where:

1. **Urban Efficiency Leaders**

- Woodsmoke Reduction (0.1027)
- Fluorinated Gases (0.0380)
- Renewable Energy for Agriculture (0.0290)
- But notice: Top performers have relatively small funding amounts (except Waste Diversion at $45M)

2. **Rural Efficiency Champions**

- Sustainable Agricultural Lands (0.1273, $108M)
- Low-Carbon Fuels (0.0905)
- Larger programs achieving high efficiency
- Strong agriculture and land-use focus

3. **Central Valley Success Stories**

- Sustainable Agricultural Lands (0.1253)
- Climate Smart Agriculture (0.0751, $274M)
- High efficiency with substantial funding amounts

Key Patterns: 1. Agricultural and land-use programs dominate efficiency in rural/CV 2. Urban areas achieve efficiency through smaller, targeted programs 3. Central Valley shows ability to maintain efficiency at scale 4. Similar program types (e.g., Food Production) perform differently by region

Would you like to: 1. Analyze the characteristics of the most efficient programs in each region? 2. Look at how program size relates to efficiency within each region? 3. Examine whether these patterns change over time? 4. Investigate if multi-county collaborations affect these regional differences?

[18]:
```
# Analyze multi-county projects by region type
data_filtered['is_multi_county'] = data_filtered['County'].str.contains(',',␣
 ↪na=False)

regional_collab = data_filtered.groupby(['region_type', 'is_multi_county']).
 ↪agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Total Project GHGReductions': 'sum',
    'Is Benefit Disadvantaged Communities': 'mean'
}).round(3)

# Calculate efficiency
regional_collab['ghg_per_dollar'] = (
    regional_collab[('Total Project GHGReductions', 'sum')] /
    regional_collab[('Total Program GGRFFunding', 'sum')]
)

print("Multi-County vs Single-County by Region:")
print(regional_collab)
```

```
Multi-County vs Single-County by Region:
                             Total Program GGRFFunding              \
                                                 count         sum
region_type    is_multi_county
Central Valley False                             14312  1314145988
               True                                184    73786209
Rural          False                              8194  2060718991
               True                                280   122681272
Urban          False                              7893  2264345480
               True                                288   324607138

                                      Total Project GHGReductions  \
                                                              mean
region_type    is_multi_county
Central Valley False                                    91821.268
               True                                    401012.005
Rural          False                                    251491.212
               True                                    438147.400
Urban          False                                    286880.208
               True                                   1127108.118

                                                  sum
region_type    is_multi_county
Central Valley False                         24949448
               True                            371193
Rural          False                         22645122
               True                            906509
Urban          False                         12069861
               True                           4945575

                             Is Benefit Disadvantaged Communities  \
                                                             mean
region_type    is_multi_county
Central Valley False                                        0.182
               True                                         0.103
Rural          False                                        0.238
               True                                         0.086
```

```
Urban          False                                          0.474
               True                                           0.135

                                ghg_per_dollar

region_type    is_multi_county
Central Valley False                0.018985
               True                 0.005031
Rural          False                0.010989
               True                 0.007389
Urban          False                0.005330
               True                 0.015236
```

This intersection of multi-county and regional patterns reveals:

1. **Project Size Patterns**

- Average project size is consistently larger in multi-county projects across all regions:
  - Urban: $1.13M vs $287K
  - Rural: $438K vs $251K
  - Central Valley: $401K vs $92K

1. **GHG Efficiency Varies By Region**

- Single-county efficiency:
  - Central Valley leads (0.019)
  - Rural second (0.011)
  - Urban lowest (0.005)
- Multi-county efficiency shows different pattern:
  - Urban leads (0.015)
  - Rural and Central Valley lower (0.007 and 0.005)

1. **DAC Benefits Trade-off**

- Single-county DAC rates:
  - Urban highest (47.4%)
  - Rural (23.8%)
  - Central Valley (18.2%)
- Multi-county shows lower DAC rates across all regions:
  - Urban (13.5%)
  - Central Valley (10.3%)
  - Rural lowest (8.6%)

Key Findings: 1. Multi-county collaborations most successful in urban areas for GHG efficiency 2. Single-county projects better at reaching DACs across all regions 3. Central Valley efficiency advantage disappears in multi-county projects 4. Urban areas seem to benefit most from regional coordination

This suggests different collaboration strategies might be needed for different regions.

[19]:
```
gdf.columns
```

[19]:
```
Index(['Tract', 'ZIP', 'County', 'ApproxLoc', 'TotPop19', 'CIscore',
       'CIscoreP', 'Ozone', 'OzoneP', 'PM2_5', 'PM2_5_P', 'DieselPM',
       'DieselPM_P', 'Pesticide', 'PesticideP', 'Tox_Rel', 'Tox_Rel_P',
       'Traffic', 'TrafficP', 'DrinkWat', 'DrinkWatP', 'Lead', 'Lead_P',
       'Cleanup', 'CleanupP', 'GWThreat', 'GWThreatP', 'HazWaste', 'HazWasteP',
       'ImpWatBod', 'ImpWatBodP', 'SolWaste', 'SolWasteP', 'PollBurd',
       'PolBurdSc', 'PolBurdP', 'Asthma', 'AsthmaP', 'LowBirtWt', 'LowBirWP',
       'Cardiovas', 'CardiovasP', 'Educatn', 'EducatP', 'Ling_Isol',
       'Ling_IsolP', 'Poverty', 'PovertyP', 'Unempl', 'UnemplP', 'HousBurd',
       'HousBurdP', 'PopChar', 'PopCharSc', 'PopCharP', 'Child_10',
       'Pop_10_64', 'Elderly65', 'Hispanic', 'White', 'AfricanAm', 'NativeAm',
       'OtherMult', 'Shape_Leng', 'Shape_Area', 'AAPI', 'geometry'],
      dtype='object')
```

[20]:
```
# Clean census tract IDs in both dataframes
def standardize_tract(tract):
    if pd.isna(tract):
        return None
    if isinstance(tract, (int, float)):
        return str(int(tract))
    return str(tract).strip()

# Check the data first
print("\nCES Tract Examples:")
print(gdf['Tract'].head())
print("\nCCI Tract Examples:")
print(data_filtered['Census Tract'].head())

# Clean and standardize
gdf['Tract_clean'] = gdf['Tract'].apply(standardize_tract)
data_filtered['Tract_clean'] = data_filtered['Census Tract'].
 ↪apply(standardize_tract)

# Check for missing tracts
print("\nMissing Tract Counts:")
print("CES missing tracts:", gdf['Tract_clean'].isna().sum())
print("CCI missing tracts:", data_filtered['Tract_clean'].isna().sum())

# Merge datasets, excluding null tracts
merged_data = pd.merge(
    data_filtered[data_filtered['Tract_clean'].notna()],
    gdf[['Tract_clean', 'CIscore', 'CIscoreP', 'PollBurd', 'Poverty',␣
 ↪'TotPop19']],
    on='Tract_clean',
    how='inner'
)
```

```
print("\nMerged Data Summary:")
print("Total rows:", len(merged_data))
print("Unique tracts:", merged_data['Tract_clean'].nunique())

# Basic analysis of funding by CES score
print("\nFunding Distribution by CalEnviroScreen Score Quintiles:")
merged_data['CES_quintile'] = pd.qcut(merged_data['CIscore'], q=5,
                                      labels=['Lowest Burden', 'Low', 'Medium',␣
 ↪'High', 'Highest Burden'])

ces_analysis = merged_data.groupby('CES_quintile').agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Total Project GHGReductions': 'sum',
    'Is Benefit Disadvantaged Communities': 'mean'
}).round(2)

print(ces_analysis)
```

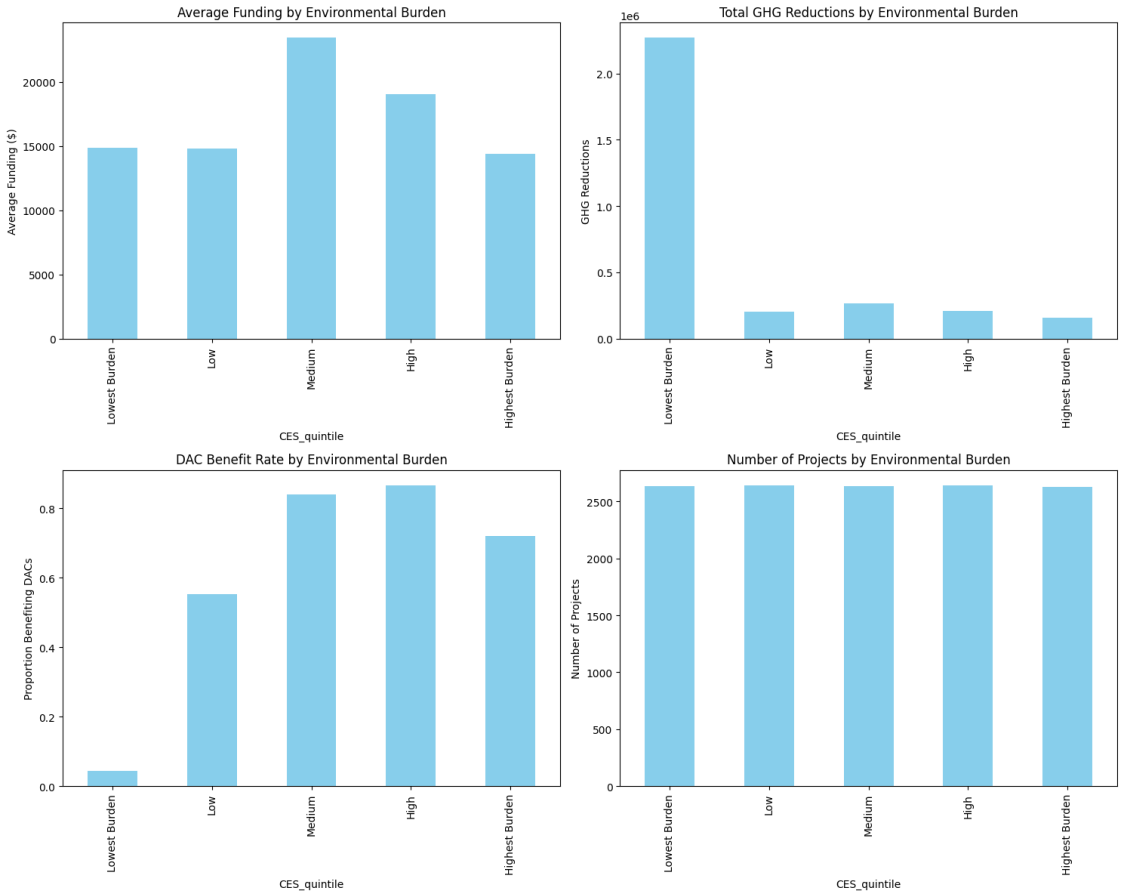
```
CES Tract Examples:
0    6.083002e+09
1    6.083002e+09
2    6.083002e+09
3    6.083002e+09
4    6.083002e+09
Name: Tract, dtype: float64

CCI Tract Examples:
16067             NaN
25098    6.001402e+09
25099    6.001402e+09
25100    6.001406e+09
25101    6.001406e+09
Name: Census Tract, dtype: float64

Missing Tract Counts:
CES missing tracts: 0
CCI missing tracts: 17967

Merged Data Summary:
Total rows: 13184
Unique tracts: 3665

Funding Distribution by CalEnviroScreen Score Quintiles:
               Total Program GGRFFunding                       \
                                   count       sum      mean
CES_quintile
Lowest Burden                       2638  39278353  14889.44
```

```
Low                                 2642  39059436  14784.04
Medium                              2636  61809955  23448.39
High                                2643  50317266  19037.94
Highest Burden                      2625  37787591  14395.27

               Total Project GHGReductions  \
                                        sum
CES_quintile
Lowest Burden                       2270359
Low                                  202457
Medium                               265818
High                                 210633
Highest Burden                       159776

               Is Benefit Disadvantaged Communities
                                               mean
CES_quintile
Lowest Burden                                  0.04
Low                                            0.55
Medium                                         0.84
High                                           0.87
Highest Burden                                 0.72

/tmp/ipykernel_164030/1664997265.py:41: FutureWarning: The default of
observed=False is deprecated and will be changed to True in a future version of
pandas. Pass observed=False to retain current behavior or observed=True to adopt
the future default and silence this warning.
  ces_analysis = merged_data.groupby('CES_quintile').agg({
```

[21]:
```
import matplotlib.pyplot as plt

fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(15, 12))

# Average Funding by CES Score
merged_data.groupby('CES_quintile')['Total Program GGRFFunding'].mean().plot(
    kind='bar', ax=ax1, color='skyblue')
ax1.set_title('Average Funding by Environmental Burden')
ax1.set_ylabel('Average Funding ($)')

# GHG Reductions
merged_data.groupby('CES_quintile')['Total Project GHGReductions'].sum().plot(
    kind='bar', ax=ax2, color='skyblue')
ax2.set_title('Total GHG Reductions by Environmental Burden')
ax2.set_ylabel('GHG Reductions')

# DAC Benefit Rate
merged_data.groupby('CES_quintile')['Is Benefit Disadvantaged Communities'].
 ↪mean().plot(
```

```
    kind='bar', ax=ax3, color='skyblue')
ax3.set_title('DAC Benefit Rate by Environmental Burden')
ax3.set_ylabel('Proportion Benefiting DACs')

# Project Counts
merged_data.groupby('CES_quintile')['Total Program GGRFFunding'].count().plot(
    kind='bar', ax=ax4, color='skyblue')
ax4.set_title('Number of Projects by Environmental Burden')
ax4.set_ylabel('Number of Projects')

plt.tight_layout()
plt.show()

# Also examine program types by CES quintile
print("\nProgram Types by Environmental Burden:")
program_dist = pd.crosstab(
    merged_data['CES_quintile'],
    merged_data['Program Name'],
    values=merged_data['Total Program GGRFFunding'],
    aggfunc='sum'
)/1e6  # Convert to millions

print(program_dist.round(2))
```

```
/tmp/ipykernel_164030/541290860.py:6: FutureWarning: The default of
observed=False is deprecated and will be changed to True in a future version of
pandas. Pass observed=False to retain current behavior or observed=True to adopt
the future default and silence this warning.
  merged_data.groupby('CES_quintile')['Total Program GGRFFunding'].mean().plot(
/tmp/ipykernel_164030/541290860.py:12: FutureWarning: The default of
observed=False is deprecated and will be changed to True in a future version of
pandas. Pass observed=False to retain current behavior or observed=True to adopt
the future default and silence this warning.
  merged_data.groupby('CES_quintile')['Total Project GHGReductions'].sum().plot(
/tmp/ipykernel_164030/541290860.py:18: FutureWarning: The default of
observed=False is deprecated and will be changed to True in a future version of
pandas. Pass observed=False to retain current behavior or observed=True to adopt
the future default and silence this warning.
  merged_data.groupby('CES_quintile')['Is Benefit Disadvantaged
Communities'].mean().plot(
/tmp/ipykernel_164030/541290860.py:24: FutureWarning: The default of
observed=False is deprecated and will be changed to True in a future version of
pandas. Pass observed=False to retain current behavior or observed=True to adopt
the future default and silence this warning.
  merged_data.groupby('CES_quintile')['Total Program GGRFFunding'].count().plot(
```



Program Types by Environmental Burden:

Program Name	Low-Income Weatherization Program	\
CES_quintile		
Lowest Burden	3.69	
Low	28.77	
Medium	50.45	
High	45.19	
Highest Burden	30.88	

Program Name	Sustainable Agricultural Lands Conservation Program	\
CES_quintile		
Lowest Burden	22.05	
Low	0.00	
Medium	0.00	
High	0.00	
Highest Burden	0.00	

Program Name	Transformative Climate Communities	Waste Diversion	\
CES_quintile			

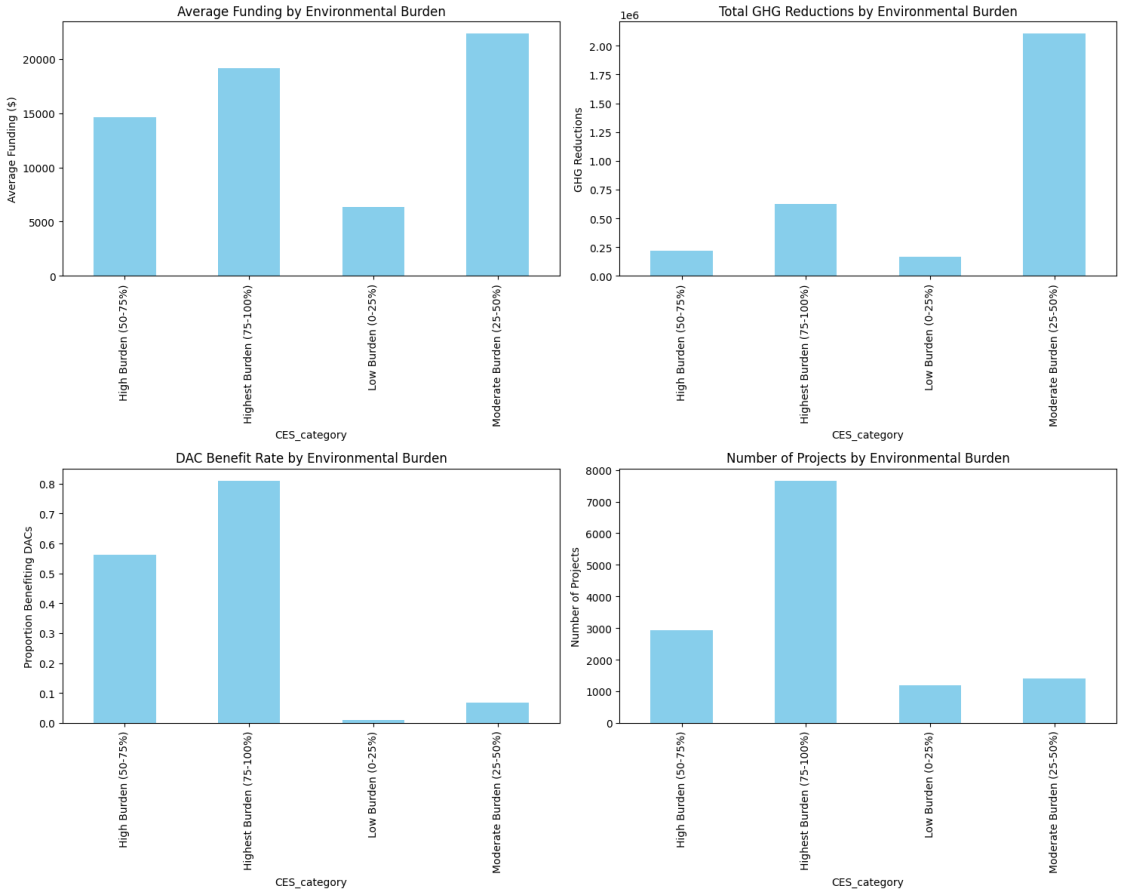
```
Lowest Burden                                0.00                0.34
Low                                          0.48                0.23
Medium                                       0.94                0.15
High                                         0.33                0.12
Highest Burden                               3.14                0.14

Program Name    Water-Energy Efficiency  Woodsmoke Reduction Program  \
CES_quintile
Lowest Burden                      7.13                         6.00
Low                                8.02                         1.50
Medium                             9.63                         0.36
High                               4.31                         0.01
Highest Burden                     3.38                         0.03

Program Name    Workforce Training and Development
CES_quintile
Lowest Burden                                0.07
Low                                          0.07
Medium                                       0.29
High                                         0.36
Highest Burden                               0.21
```

[22]:
```python
# Create meaningful CES score categories
def categorize_ces(score):
    if score >= 75:  # Top 25% - most burdened
        return 'Highest Burden (75-100%)'
    elif score >= 50:
        return 'High Burden (50-75%)'
    elif score >= 25:
        return 'Moderate Burden (25-50%)'
    else:  # Bottom 25% - least burdened
        return 'Low Burden (0-25%)'

# Add new categorization
merged_data['CES_category'] = merged_data['CIscoreP'].apply(categorize_ces)

# Analyze distribution with new categories
ces_analysis_new = merged_data.groupby('CES_category').agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Total Project GHGReductions': 'sum',
    'Is Benefit Disadvantaged Communities': 'mean'
}).round(2)

print("Distribution using CES Percentile Thresholds:")
print(ces_analysis_new)

# Visualize with new categories
```

```
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(15, 12))

# Average Funding
merged_data.groupby('CES_category')['Total Program GGRFFunding'].mean().plot(
    kind='bar', ax=ax1, color='skyblue')
ax1.set_title('Average Funding by Environmental Burden')
ax1.set_ylabel('Average Funding ($)')

# GHG Reductions
merged_data.groupby('CES_category')['Total Project GHGReductions'].sum().plot(
    kind='bar', ax=ax2, color='skyblue')
ax2.set_title('Total GHG Reductions by Environmental Burden')
ax2.set_ylabel('GHG Reductions')

# DAC Benefit Rate
merged_data.groupby('CES_category')['Is Benefit Disadvantaged Communities'].
 ↪mean().plot(
    kind='bar', ax=ax3, color='skyblue')
ax3.set_title('DAC Benefit Rate by Environmental Burden')
ax3.set_ylabel('Proportion Benefiting DACs')

# Project Counts
merged_data.groupby('CES_category')['Total Program GGRFFunding'].count().plot(
    kind='bar', ax=ax4, color='skyblue')
ax4.set_title('Number of Projects by Environmental Burden')
ax4.set_ylabel('Number of Projects')

plt.tight_layout()
plt.show()
```

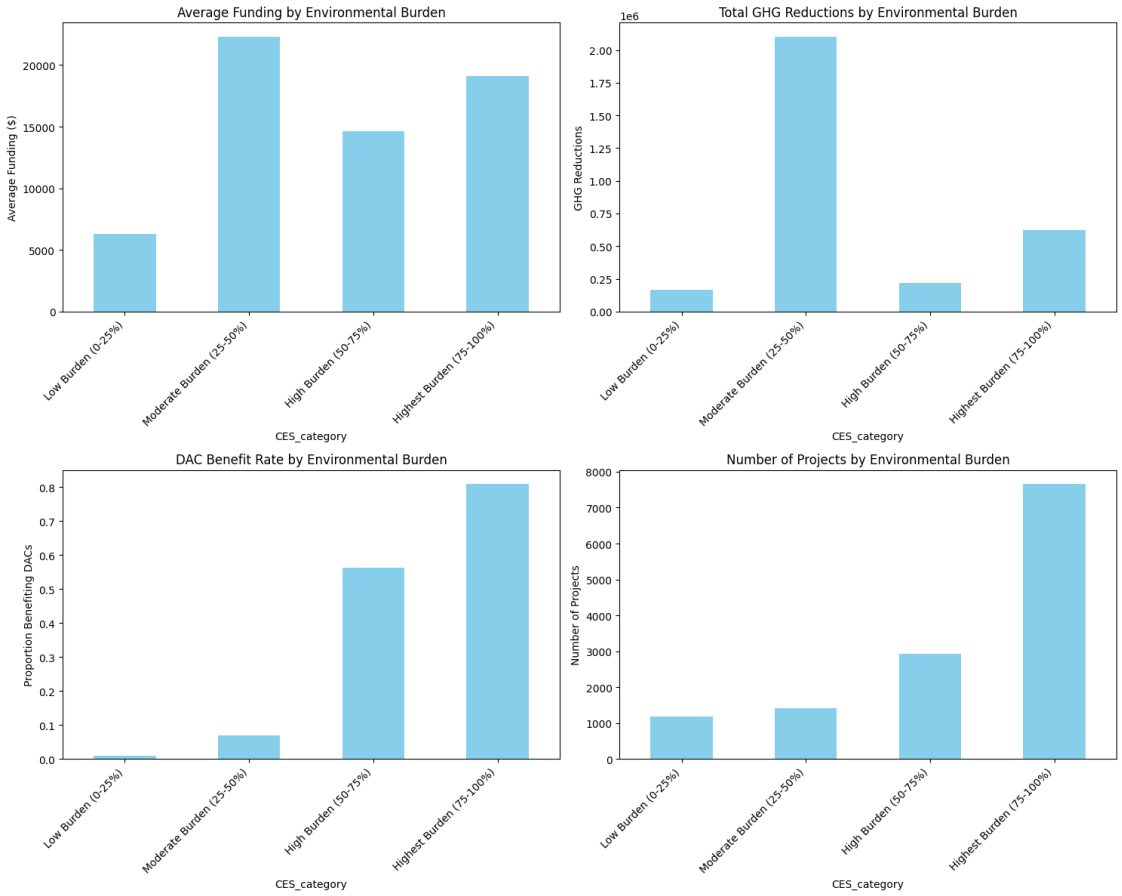
```
Distribution using CES Percentile Thresholds:
                        Total Program GGRFFunding                       \
                                            count        sum      mean
CES_category
High Burden (50-75%)                         2932   42831569  14608.31
Highest Burden (75-100%)                     7661  146493459  19121.98
Low Burden (0-25%)                           1182    7475213   6324.21
Moderate Burden (25-50%)                     1409   31452360  22322.47

                        Total Project GHGReductions  \
                                                sum
CES_category
High Burden (50-75%)                         218725
Highest Burden (75-100%)                     622363
Low Burden (0-25%)                           163482
Moderate Burden (25-50%)                    2104473

                        Is Benefit Disadvantaged Communities
```

CES_category	mean
High Burden (50-75%)	0.56
Highest Burden (75-100%)	0.81
Low Burden (0-25%)	0.01
Moderate Burden (25-50%)	0.07



[23]:
```
# Define category order
category_order = [
    'Low Burden (0-25%)',
    'Moderate Burden (25-50%)',
    'High Burden (50-75%)',
    'Highest Burden (75-100%)'
]

# Create visualizations with ordered categories
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(15, 12))

# Average Funding
```

```
merged_data.groupby('CES_category')['Total Program GGRFFunding'].
  ↪mean()[category_order].plot(
    kind='bar', ax=ax1, color='skyblue')
ax1.set_title('Average Funding by Environmental Burden')
ax1.set_ylabel('Average Funding ($)')

# GHG Reductions
merged_data.groupby('CES_category')['Total Project GHGReductions'].
  ↪sum()[category_order].plot(
    kind='bar', ax=ax2, color='skyblue')
ax2.set_title('Total GHG Reductions by Environmental Burden')
ax2.set_ylabel('GHG Reductions')

# DAC Benefit Rate
merged_data.groupby('CES_category')['Is Benefit Disadvantaged Communities'].
  ↪mean()[category_order].plot(
    kind='bar', ax=ax3, color='skyblue')
ax3.set_title('DAC Benefit Rate by Environmental Burden')
ax3.set_ylabel('Proportion Benefiting DACs')

# Project Counts
merged_data.groupby('CES_category')['Total Program GGRFFunding'].
  ↪count()[category_order].plot(
    kind='bar', ax=ax4, color='skyblue')
ax4.set_title('Number of Projects by Environmental Burden')
ax4.set_ylabel('Number of Projects')

# Rotate x-labels for better readability
for ax in [ax1, ax2, ax3, ax4]:
    ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha='right')

plt.tight_layout()
plt.show()
```



[24]:
```
# Agency Performance Analysis
agency_performance = data_filtered.groupby('Agency Name').agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Total Project GHGReductions': 'sum',
    'Is Benefit Disadvantaged Communities': 'mean'
}).round(3)

print("Agency Performance Overview:")
print(agency_performance)

# Add efficiency metrics
agency_performance['ghg_per_dollar'] = (
    agency_performance[('Total Project GHGReductions', 'sum')] /
    agency_performance[('Total Program GGRFFunding', 'sum')]
)

# Look at agency patterns in multi-county projects
multi_county = data_filtered[data_filtered['County'].str.contains(',',␣
 ↪na=False)]
```

```python
multi_agency_patterns = multi_county.groupby('Agency Name').agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Total Project GHGReductions': 'sum',
    'Is Benefit Disadvantaged Communities': 'mean'
}).round(3)

print("\nAgency Performance in Multi-County Projects:")
print(multi_agency_patterns)

# Visualize agency performance metrics
plt.figure(figsize=(15, 10))

# Plot GHG efficiency vs DAC benefit rate by agency
plt.scatter(
    agency_performance['ghg_per_dollar'],
    agency_performance[('Is Benefit Disadvantaged Communities', 'mean')],
    s=agency_performance[('Total Program GGRFFunding', 'sum')]/1e6,  # Size by␣
 ↪total funding
    alpha=0.6
)

# Add agency labels
for idx, row in agency_performance.iterrows():
    plt.annotate(idx[:20] + '...' if len(idx) > 20 else idx,
                 (row['ghg_per_dollar'],
                  row[('Is Benefit Disadvantaged Communities', 'mean')]),
                 xytext=(5, 5), textcoords='offset points')

plt.xlabel('GHG Reduction per Dollar')
plt.ylabel('DAC Benefit Rate')
plt.title('Agency Performance: Climate Impact vs Equity\nSize = Total Funding')

plt.tight_layout()
plt.show()
```

```
Agency Performance Overview:
                                                Total Program GGRFFunding  \
                                                                    count
Agency Name
California Air Resources Board                                      12329
California Coastal Commission                                          17
California Conservation Corps                                         766
California Department of Community Services and…                     6331
California Department of Fish and Wildlife                              8
California Department of Food and Agriculture                        1311
California Department of Forestry and Fire Prot…                     2047
California Department of Resources Recycling an…                      723
California Department of Transportation                               769
```

California Department of Water Resources	4887
California Energy Commission	119
California Environmental Protection Agency	2
California Governors Office of Emergency Services	63
California Natural Resources Agency	213
California State Coastal Conservancy	10
California State Transportation Agency	135
California State Water Resources Control Board	100
California Strategic Growth Council	1200
California Wildlife Conservation Board	17
California Workforce Development Board	30
San Francisco Bay Conservation and Development …	74

\

Agency Name	sum	mean
California Air Resources Board	796368463	64593.111
California Coastal Commission	2667413	156906.647
California Conservation Corps	47161363	61568.359
California Department of Community Services and…	158967758	25109.423
California Department of Fish and Wildlife	16171301	2021412.625
California Department of Food and Agriculture	338161549	257941.685
California Department of Forestry and Fire Prot…	967562562	472673.455
California Department of Resources Recycling an…	175129441	242226.059
California Department of Transportation	785906434	1021984.960
California Department of Water Resources	52470432	10736.737
California Energy Commission	140791478	1183121.664
California Environmental Protection Agency	2560940	1280470.000
California Governors Office of Emergency Services	31037358	492656.476
California Natural Resources Agency	145023190	680860.047
California State Coastal Conservancy	7218244	721824.400
California State Transportation Agency	771556000	5715229.630
California State Water Resources Control Board	118601405	1186014.050
California Strategic Growth Council	1557246299	1297705.249
California Wildlife Conservation Board	12574800	739694.118
California Workforce Development Board	24428357	814278.567
San Francisco Bay Conservation and Development …	8680291	117301.230

Total Project GHGReductions

\

Agency Name	sum
California Air Resources Board	509888
California Coastal Commission	0
California Conservation Corps	292164
California Department of Community Services and…	583978
California Department of Fish and Wildlife	405413
California Department of Food and Agriculture	21019827

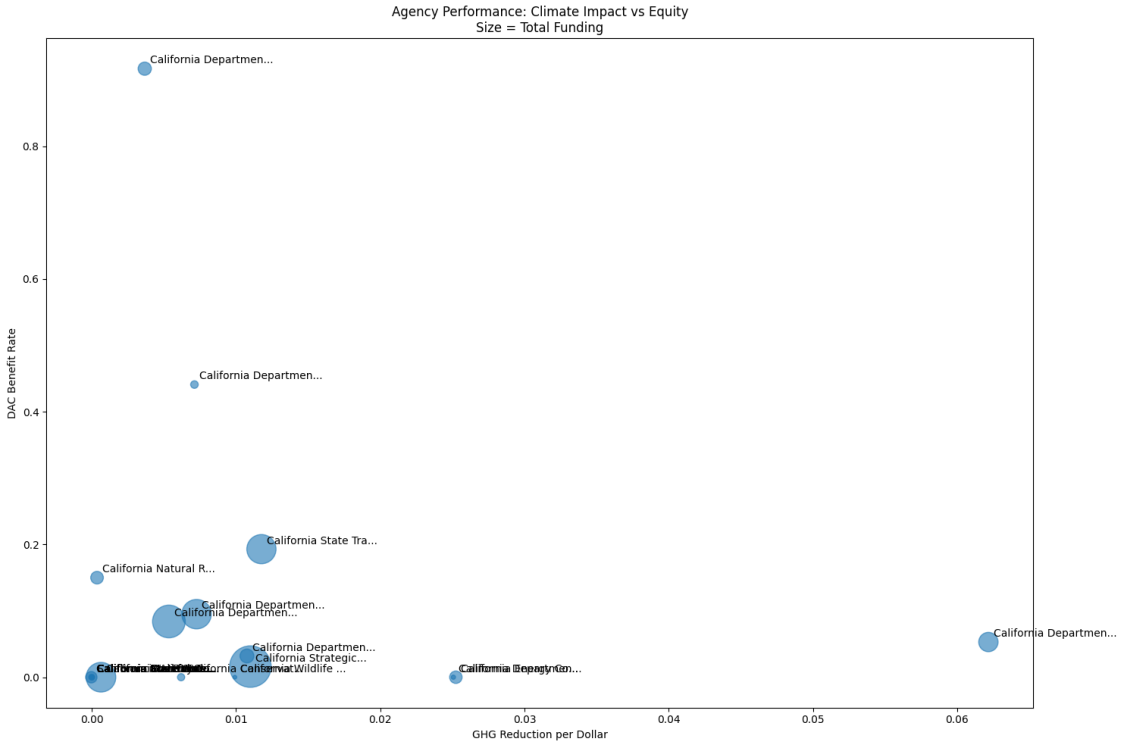
```
California Department of Forestry and Fire Prot…                    5176685
California Department of Resources Recycling an…                    1884352
California Department of Transportation                               5709880
California Department of Water Resources                               373494
California Energy Commission                                          3553812
California Environmental Protection Agency                                  0
California Governors Office of Emergency Services                           0
California Natural Resources Agency                                    53327
California State Coastal Conservancy                                       0
California State Transportation Agency                                9076036
California State Water Resources Control Board                          -3620
California Strategic Growth Council                                  17127711
California Wildlife Conservation Board                                 124761
California Workforce Development Board                                      0
San Francisco Bay Conservation and Development …                            0
```

```
                                                    Is Benefit Disadvantaged
Communities
   mean
Agency Name
California Air Resources Board
0.000
California Coastal Commission
0.000
California Conservation Corps
0.000
California Department of Community Services and…
0.917
California Department of Fish and Wildlife
0.000
California Department of Food and Agriculture
0.053
California Department of Forestry and Fire Prot…
0.084
California Department of Resources Recycling an…
0.032
California Department of Transportation
0.095
California Department of Water Resources
0.441
California Energy Commission
0.000
California Environmental Protection Agency
0.000
California Governors Office of Emergency Services
0.000
California Natural Resources Agency
0.150
```

```
California State Coastal Conservancy
0.000
California State Transportation Agency
0.193
California State Water Resources Control Board
0.000
California Strategic Growth Council
0.016
California Wildlife Conservation Board
0.000
California Workforce Development Board
0.000
San Francisco Bay Conservation and Development …
0.000

Agency Performance in Multi-County Projects:
                                                Total Program GGRFFunding  \
                                                                    count
Agency Name
California Air Resources Board                                         25
California Conservation Corps                                           9
California Department of Food and Agriculture                          41
California Department of Forestry and Fire Prot…                      339
California Department of Resources Recycling an…                       23
California Department of Transportation                               158
California Energy Commission                                            9
California Natural Resources Agency                                     7
California State Transportation Agency                                 57
California State Water Resources Control Board                         45
California Strategic Growth Council                                    32
California Wildlife Conservation Board                                  7

                                                                        \
                                                       sum          mean
Agency Name
California Air Resources Board                     2104048    84161.920
California Conservation Corps                      1248353   138705.889
California Department of Food and Agriculture      6296871   153582.220
California Department of Forestry and Fire Prot…  32351508    95432.177
California Department of Resources Recycling an…   6962204   302704.522
California Department of Transportation          162505329  1028514.741
California Energy Commission                      12379666  1375518.444
California Natural Resources Agency                2879148   411306.857
California State Transportation Agency           232590000  4080526.316
California State Water Resources Control Board    46265001  1028111.133
California Strategic Growth Council               11948291   373384.094
California Wildlife Conservation Board             3544200   506314.286
```

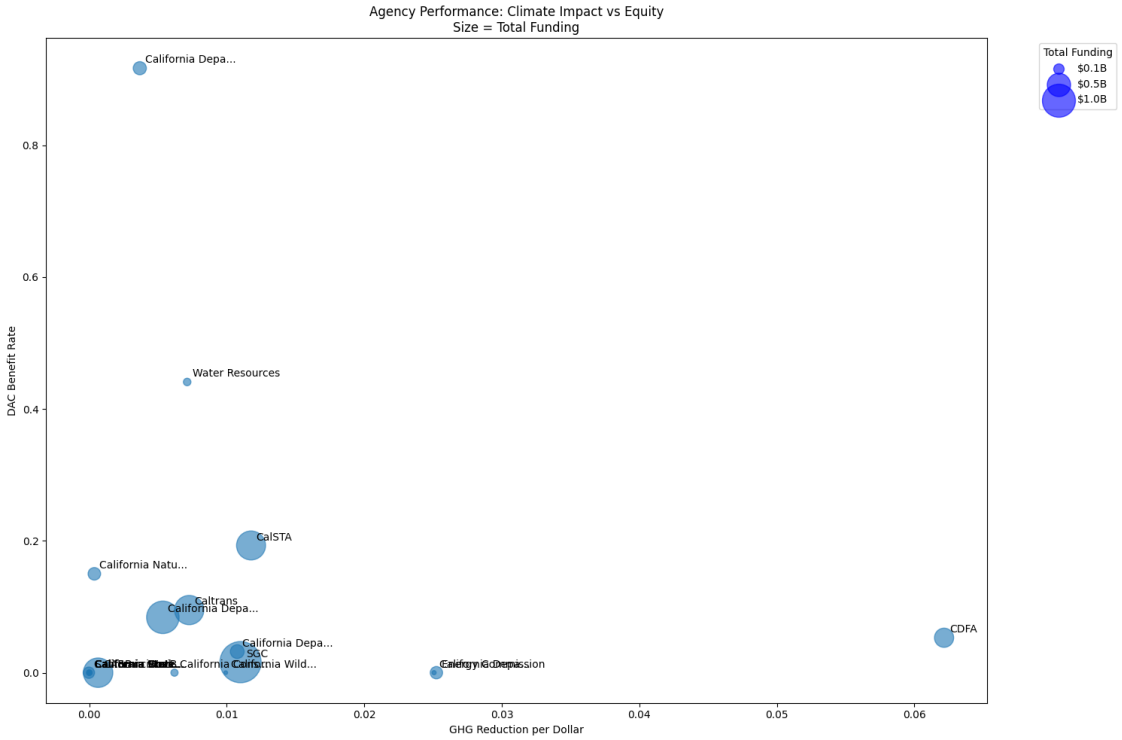
```
                                                    Total Project GHGReductions
\
                                                                            sum
Agency Name
California Air Resources Board                                             -413
California Conservation Corps                                               842
California Department of Food and Agriculture                             42951
California Department of Forestry and Fire Prot…                         185424
California Department of Resources Recycling an…                          35434
California Department of Transportation                                 2460280
California Energy Commission                                             501007
California Natural Resources Agency                                        1191
California State Transportation Agency                                  2748646
California State Water Resources Control Board                            -3234
California Strategic Growth Council                                      251149
California Wildlife Conservation Board                                        0

                                                    Is Benefit Disadvantaged
Communities
   mean
Agency Name
California Air Resources Board
0.000
California Conservation Corps
0.000
California Department of Food and Agriculture
0.000
California Department of Forestry and Fire Prot…
0.168
California Department of Resources Recycling an…
0.130
California Department of Transportation
0.089
California Energy Commission
0.000
California Natural Resources Agency
0.000
California State Transportation Agency
0.105
California State Water Resources Control Board
0.000
California Strategic Growth Council
0.062
California Wildlife Conservation Board
0.000

/home/dadams/Repos/california_equity_git/.venv/lib/python3.12/site-
packages/matplotlib/text.py:1465: FutureWarning: Calling float on a single
```

```
element Series is deprecated and will raise a TypeError in the future. Use
float(ser.iloc[0]) instead
  x = float(self.convert_xunits(x))
```



[25]:
```
plt.figure(figsize=(15, 10))

# Create shorter but still meaningful agency names
agency_shortnames = {
    'California Air Resources Board': 'CARB',
    'California Department of Community Services and...': 'Community Services',
    'California Department of Food and Agriculture': 'CDFA',
    'California Department of Forestry and Fire Prot...': 'CalFire',
    'California Department of Resources Recycling an...': 'CalRecycle',
    'California Department of Transportation': 'Caltrans',
    'California Department of Water Resources': 'Water Resources',
    'California Energy Commission': 'Energy Commission',
    'California State Transportation Agency': 'CalSTA',
    'California Strategic Growth Council': 'SGC',
    # Add other agencies as needed
}

# Plot GHG efficiency vs DAC benefit rate by agency
plt.scatter(
```

```
    agency_performance['ghg_per_dollar'],
    agency_performance[('Is Benefit Disadvantaged Communities', 'mean')],
    s=agency_performance[('Total Program GGRFFunding', 'sum')]/1e6,  # Size by␣
 ↪total funding
    alpha=0.6
)

# Add agency labels with shorter names
for idx, row in agency_performance.iterrows():
    short_name = agency_shortnames.get(idx, idx[:15] + '...')
    plt.annotate(short_name,
                 (row['ghg_per_dollar'],
                  row[('Is Benefit Disadvantaged Communities', 'mean')]),
                 xytext=(5, 5), textcoords='offset points')

plt.xlabel('GHG Reduction per Dollar')
plt.ylabel('DAC Benefit Rate')
plt.title('Agency Performance: Climate Impact vs Equity\nSize = Total Funding')

# Add legend for bubble size
legend_elements = [plt.scatter([], [], s=s, label=f'${l}B', alpha=0.6,␣
 ↪color='blue')
                   for s, l in zip([100, 500, 1000], ['0.1', '0.5', '1.0'])]
plt.legend(handles=legend_elements, title='Total Funding',
           bbox_to_anchor=(1.05, 1), loc='upper left')

plt.tight_layout()
plt.show()
```



1. **Strong GHG Performers**

- California Department of Food and Agriculture shows highest GHG efficiency (~0.06 reductions per dollar)
- California Energy Commission also performs well (~0.03)
- However, both show low DAC benefit rates (0-5%)

1. **Strong Equity Performers**

- California Department of Community Services leads in DAC benefits (91.7%)
- Department of Water Resources shows good equity performance (44.1%)
- But these have lower GHG efficiency

1. **Large Scale Implementers** (shown by bubble size)

- Strategic Growth Council ($1.56B)
- CalSTA ($772M)
- CARB ($796M)
- Department of Forestry and Fire Protection ($968M)

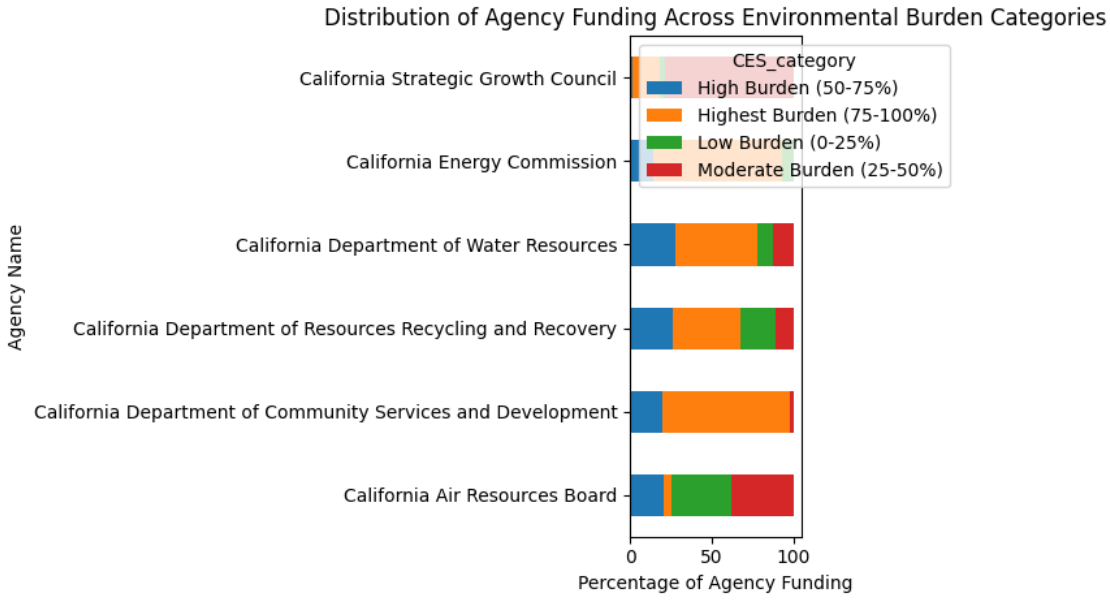
1. **Multi-County Implementation**

- Transportation agencies lead in multi-county projects (CalSTA: 57 projects, $233M)
- CalFire shows significant multi-county work (339 projects)
- Strategic Growth Council's multi-county projects show lower DAC rates than their overall portfolio

[26]:
```
# Analyze agency performance across CES score categories
merged_agency_ces = merged_data.groupby(['Agency Name', 'CES_category']).agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Total Project GHGReductions': 'sum',
    'Is Benefit Disadvantaged Communities': 'mean'
}).round(3)

# Calculate percentage of agency funding going to each CES category
agency_funding = pd.pivot_table(
    merged_data,
    values='Total Program GGRFFunding',
    index='Agency Name',
    columns='CES_category',
    aggfunc='sum'
).fillna(0)

# Calculate percentages
agency_funding_pct = agency_funding.div(agency_funding.sum(axis=1), axis=0) *␣
 ↪100

print("Percentage of Agency Funding by Environmental Burden Category:")
print(agency_funding_pct.round(2))

# Visualize distribution patterns
plt.figure(figsize=(15, 8))
agency_funding_pct.plot(kind='barh', stacked=True)
plt.title('Distribution of Agency Funding Across Environmental Burden␣
 ↪Categories')
plt.xlabel('Percentage of Agency Funding')
plt.tight_layout()
plt.show()

# Look at average project size by burden category for each agency
avg_project_size = merged_data.groupby(['Agency Name', 'CES_category'])['Total␣
 ↪Program GGRFFunding'].mean()/1e6
avg_project_size = avg_project_size.unstack()
print("\nAverage Project Size (Millions $) by Environmental Burden Category:")
print(avg_project_size.round(2))
```

```
Percentage of Agency Funding by Environmental Burden Category:
CES_category                                        High Burden (50-75%)  \
Agency Name
California Air Resources Board                                     20.81
California Department of Community Services and…                   19.77
California Department of Resources Recycling an…                   26.02
California Department of Water Resources                           27.36
California Energy Commission                                       14.29
```

```
California Strategic Growth Council                          1.78

CES_category                                    Highest Burden (75-100%)  \
Agency Name
California Air Resources Board                                   4.92
California Department of Community Services and…                77.98
California Department of Resources Recycling an…                41.17
California Department of Water Resources                        50.91
California Energy Commission                                    78.57
California Strategic Growth Council                             16.38

CES_category                                    Low Burden (0-25%)  \
Agency Name
California Air Resources Board                           36.65
California Department of Community Services and…          0.29
California Department of Resources Recycling an…         21.83
California Department of Water Resources                  9.21
California Energy Commission                              7.14
California Strategic Growth Council                       3.15

CES_category                                    Moderate Burden (25-50%)
Agency Name
California Air Resources Board                                  37.63
California Department of Community Services and…                1.96
California Department of Resources Recycling an…               10.98
California Department of Water Resources                        12.51
California Energy Commission                                     0.00
California Strategic Growth Council                             78.69

<Figure size 1500x800 with 0 Axes>
```



```
Average Project Size (Millions $) by Environmental Burden Category:
CES_category                                        High Burden (50-75%)  \
Agency Name
California Air Resources Board                                      0.01
California Department of Community Services and…                    0.03
California Department of Resources Recycling an…                    0.01
California Department of Water Resources                            0.01
California Energy Commission                                        0.07
California Strategic Growth Council                                 0.00

CES_category                                        Highest Burden (75-100%)  \
Agency Name
California Air Resources Board                                        0.01
California Department of Community Services and…                      0.02
California Department of Resources Recycling an…                      0.01
California Department of Water Resources                              0.01
California Energy Commission                                          0.07
California Strategic Growth Council                                   0.01

CES_category                                        Low Burden (0-25%)  \
Agency Name
California Air Resources Board                                   0.01
California Department of Community Services and…                 0.07
California Department of Resources Recycling an…                 0.01
California Department of Water Resources                         0.00
California Energy Commission                                     0.07
California Strategic Growth Council                              0.85

CES_category                                        Moderate Burden (25-50%)
Agency Name
California Air Resources Board                                        0.01
California Department of Community Services and…                      0.03
California Department of Resources Recycling an…                      0.01
California Department of Water Resources                              0.00
California Energy Commission                                           NaN
California Strategic Growth Council                                   1.93
```

This data reveals fascinating patterns in how different agencies serve communities with varying environmental burdens:

1. **Strong Environmental Justice Focus**

- Community Services: 78% of funding to highest burden areas
- Energy Commission: 79% to highest burden areas
- Water Resources: 51% to highest burden areas
- CalRecycle: 41% to highest burden areas

2. **More Dispersed Distribution**

- CARB shows relatively even distribution:
  - 37.6% Moderate burden
  - 36.7% Low burden
  - 20.8% High burden
  - Only 4.9% to highest burden areas

3. **Unique Strategic Growth Council Pattern**

- Heavy focus on moderate burden areas (78.7%)
- Limited focus on highest burden areas (16.4%)
- Largest average project sizes:
  - $1.93M in moderate burden areas
  - $0.85M in low burden areas
  - Much smaller in high burden areas

4. **Project Size Variations**

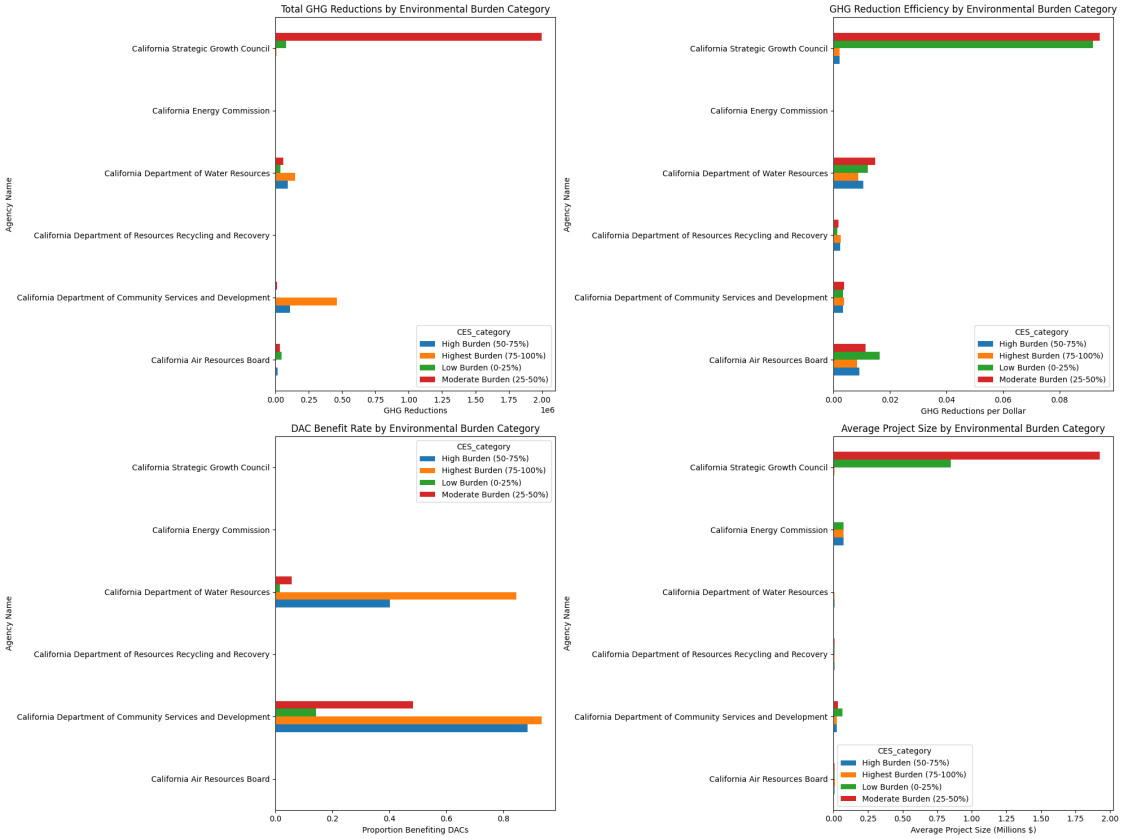
- Most agencies maintain consistent project sizes across burden categories
- Notable exceptions:
  - Strategic Growth Council: Much larger projects in moderate burden areas
  - Community Services: Larger projects in low burden areas

This suggests: 1. Different agency missions lead to different environmental justice approaches 2. Some agencies specifically target high-burden areas 3. Project size might be influenced by community capacity 4. Need for better coordination in highest burden areas

[27]:
```
# Analyze outcomes (GHG and equity) across environmental burden categories by␣
 ↪agency
outcome_analysis = merged_data.groupby(['Agency Name', 'CES_category']).agg({
   'Total Program GGRFFunding': ['count', 'sum'],
   'Total Project GHGReductions': ['sum', 'mean'],
   'Is Benefit Disadvantaged Communities': 'mean'
}).round(3)

# Calculate GHG efficiency
outcome_analysis['ghg_per_dollar'] = (
   outcome_analysis[('Total Project GHGReductions', 'sum')] /
   outcome_analysis[('Total Program GGRFFunding', 'sum')]
)

# Create visualization
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(20, 15))

# 1. Total GHG Reductions by Burden Category for each Agency
ghg_by_burden = pd.pivot_table(
   merged_data,
   values='Total Project GHGReductions',
   index='Agency Name',
```

```
    columns='CES_category',
    aggfunc='sum'
)

ghg_by_burden.plot(kind='barh', ax=ax1)
ax1.set_title('Total GHG Reductions by Environmental Burden Category')
ax1.set_xlabel('GHG Reductions')

# 2. GHG Efficiency by Burden Category
efficiency_by_burden = outcome_analysis['ghg_per_dollar'].unstack()
efficiency_by_burden.plot(kind='barh', ax=ax2)
ax2.set_title('GHG Reduction Efficiency by Environmental Burden Category')
ax2.set_xlabel('GHG Reductions per Dollar')

# 3. DAC Benefit Rate by Burden Category
dac_by_burden = pd.pivot_table(
    merged_data,
    values='Is Benefit Disadvantaged Communities',
    index='Agency Name',
    columns='CES_category',
    aggfunc='mean'
)
dac_by_burden.plot(kind='barh', ax=ax3)
ax3.set_title('DAC Benefit Rate by Environmental Burden Category')
ax3.set_xlabel('Proportion Benefiting DACs')

# 4. Average Project Size
size_by_burden = pd.pivot_table(
    merged_data,
    values='Total Program GGRFFunding',
    index='Agency Name',
    columns='CES_category',
    aggfunc='mean'
)/1e6  # Convert to millions
size_by_burden.plot(kind='barh', ax=ax4)
ax4.set_title('Average Project Size by Environmental Burden Category')
ax4.set_xlabel('Average Project Size (Millions $)')

plt.tight_layout()
plt.show()

# Print summary statistics
print("\nOutcome Analysis by Environmental Burden Category:")
print("\nGHG Efficiency (Reductions per Dollar):")
print(efficiency_by_burden.round(4))

print("\nDAC Benefit Rate:")
```

```
print(dac_by_burden.round(4))
```




```
Outcome Analysis by Environmental Burden Category:

GHG Efficiency (Reductions per Dollar):
CES_category                                        High Burden (50-75%)  \
Agency Name
California Air Resources Board                                    0.0092
California Department of Community Services and…                  0.0034
California Department of Resources Recycling an…                  0.0023
California Department of Water Resources                          0.0106
California Energy Commission                                      0.0000
California Strategic Growth Council                               0.0021

CES_category                                        Highest Burden (75-100%)  \
Agency Name
California Air Resources Board                                       0.0084
California Department of Community Services and…                     0.0037
California Department of Resources Recycling an…                     0.0026
California Department of Water Resources                             0.0088
```

```
California Energy Commission                                    0.0000
California Strategic Growth Council                             0.0022

CES_category                                  Low Burden (0-25%)  \
Agency Name
California Air Resources Board                          0.0163
California Department of Community Services and…        0.0034
California Department of Resources Recycling an…        0.0013
California Department of Water Resources                0.0122
California Energy Commission                            0.0000
California Strategic Growth Council                     0.0918

CES_category                                  Moderate Burden (25-50%)
Agency Name
California Air Resources Board                                  0.0114
California Department of Community Services and…                0.0038
California Department of Resources Recycling an…                0.0017
California Department of Water Resources                        0.0148
California Energy Commission                                       NaN
California Strategic Growth Council                             0.0943

DAC Benefit Rate:
CES_category                                  High Burden (50-75%)  \
Agency Name
California Air Resources Board                          0.0000
California Department of Community Services and…        0.8851
California Department of Resources Recycling an…        0.0000
California Department of Water Resources                0.4012
California Energy Commission                            0.0000
California Strategic Growth Council                     0.0000

CES_category                                  Highest Burden (75-100%)  \
Agency Name
California Air Resources Board                                  0.0000
California Department of Community Services and…                0.9344
California Department of Resources Recycling an…                0.0000
California Department of Water Resources                        0.8456
California Energy Commission                                    0.0000
California Strategic Growth Council                             0.0000

CES_category                                  Low Burden (0-25%)  \
Agency Name
California Air Resources Board                          0.0000
California Department of Community Services and…        0.1429
California Department of Resources Recycling an…        0.0000
California Department of Water Resources                0.0147
California Energy Commission                            0.0000
California Strategic Growth Council                     0.0000
```

```
CES_category                                    Moderate Burden (25-50%)
Agency Name
California Air Resources Board                                    0.0000
California Department of Community Services and…                  0.4835
California Department of Resources Recycling an…                  0.0000
California Department of Water Resources                          0.0565
California Energy Commission                                         NaN
California Strategic Growth Council                               0.0000
```

1. **GHG Efficiency Paradox**

- Highest efficiency in lower burden areas:
  - Strategic Growth Council: 0.0918 (low burden) vs 0.0022 (highest burden)
  - CARB: 0.0163 (low burden) vs 0.0084 (highest burden)
  - Water Resources: 0.0122 (low burden) vs 0.0088 (highest burden)
- Suggests potential infrastructure or implementation challenges in high-burden areas

2. **DAC Benefit Distribution**

- Strong targeting by some agencies:
  - Community Services: 93% DAC benefit in highest burden areas
  - Water Resources: 85% DAC benefit in highest burden areas
- Others show no DAC benefits across categories:
  - CARB
  - Energy Commission
  - Strategic Growth Council

3. **Efficiency-Equity Trade-off**

- Clear inverse relationship between GHG efficiency and DAC benefits
- Community Services shows high DAC benefits but lower GHG efficiency
- Strategic Growth Council achieves high GHG efficiency but low DAC benefits

4. **Agency Performance Patterns**

- Community Services most consistent in DAC benefits across burden levels
- Water Resources shows good balance between efficiency and equity
- Strategic Growth Council most efficient but least equitable

Key Policy Implications: 1. Need for targeted capacity building in high-burden areas 2. Potential for program design improvements to balance outcomes 3. Opportunity for cross-agency learning on successful approaches

[ ]:

[28]:
```
# 1. First, let's analyze multi-county collaborations
multi_county = data_filtered[data_filtered['County'].str.contains(',',␣
↪na=False)]
single_county = data_filtered[~data_filtered['County'].str.contains(',',␣
↪na=False)]
```

```
# Number of partner counties
multi_county['num_partners'] = multi_county['County'].str.count(',') + 1

# Basic collaboration metrics
print("Collaboration Overview:")
print(f"Number of multi-county projects: {len(multi_county)}")
print(f"Number of single-county projects: {len(single_county)}")
print("\nAverage project characteristics:")
metrics = ['Total Program GGRFFunding', 'Total Project GHGReductions', 'Is␣
 ↪Benefit Disadvantaged Communities']
print(pd.DataFrame({
    'Multi-County': multi_county[metrics].mean(),
    'Single-County': single_county[metrics].mean()
}).round(3))

# 2. Agency collaboration patterns
print("\nAgency Participation in Multi-County Projects:")
agency_collab = multi_county.groupby('Agency Name').agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Total Project GHGReductions': 'sum',
    'Is Benefit Disadvantaged Communities': 'mean',
    'num_partners': 'mean'
}).round(3)
print(agency_collab)

# 3. Visualize collaboration patterns
plt.figure(figsize=(15, 8))
plt.scatter(
    agency_collab[('num_partners', 'mean')],
    agency_collab[('Total Project GHGReductions', 'sum')] /␣
 ↪agency_collab[('Total Program GGRFFunding', 'sum')],
    s=agency_collab[('Total Program GGRFFunding', 'sum')]/1e6,
    alpha=0.6
)

# Add agency labels
for idx, row in agency_collab.iterrows():
    plt.annotate(idx[:20] + '...' if len(idx) > 20 else idx,
                 (row[('num_partners', 'mean')],
                  row[('Total Project GHGReductions', 'sum')] / row[('Total␣
 ↪Program GGRFFunding', 'sum')]),
                 xytext=(5, 5), textcoords='offset points')

plt.xlabel('Average Number of Partners')
plt.ylabel('GHG Efficiency (Reductions per Dollar)')
plt.title('Collaboration Scale vs. Efficiency\nBubble size = Total Funding')
plt.tight_layout()
```

```
plt.show()
```

```
Collaboration Overview:
Number of multi-county projects: 752
Number of single-county projects: 30399

Average project characteristics:
                                    Multi-County  Single-County
Total Program GGRFFunding               692918.376     185506.446
Total Project GHGReductions               8275.634       1962.710
Is Benefit Disadvantaged Communities         0.109          0.273

Agency Participation in Multi-County Projects:
                                                  Total Program GGRFFunding  \
                                                                      count
Agency Name
California Air Resources Board                                         25
California Conservation Corps                                           9
California Department of Food and Agriculture                          41
California Department of Forestry and Fire Prot…                      339
California Department of Resources Recycling an…                       23
California Department of Transportation                                158
California Energy Commission                                            9
California Natural Resources Agency                                     7
California State Transportation Agency                                 57
California State Water Resources Control Board                         45
California Strategic Growth Council                                    32
California Wildlife Conservation Board                                  7

                                                                        \
                                                        sum          mean
Agency Name
California Air Resources Board                      2104048    84161.920
California Conservation Corps                       1248353   138705.889
California Department of Food and Agriculture       6296871   153582.220
California Department of Forestry and Fire Prot…   32351508    95432.177
California Department of Resources Recycling an…    6962204   302704.522
California Department of Transportation           162505329  1028514.741
California Energy Commission                       12379666  1375518.444
California Natural Resources Agency                 2879148   411306.857
California State Transportation Agency            232590000  4080526.316
California State Water Resources Control Board     46265001  1028111.133
California Strategic Growth Council                11948291   373384.094
California Wildlife Conservation Board              3544200   506314.286

                                                  Total Project GHGReductions
\
                                                                         sum
```

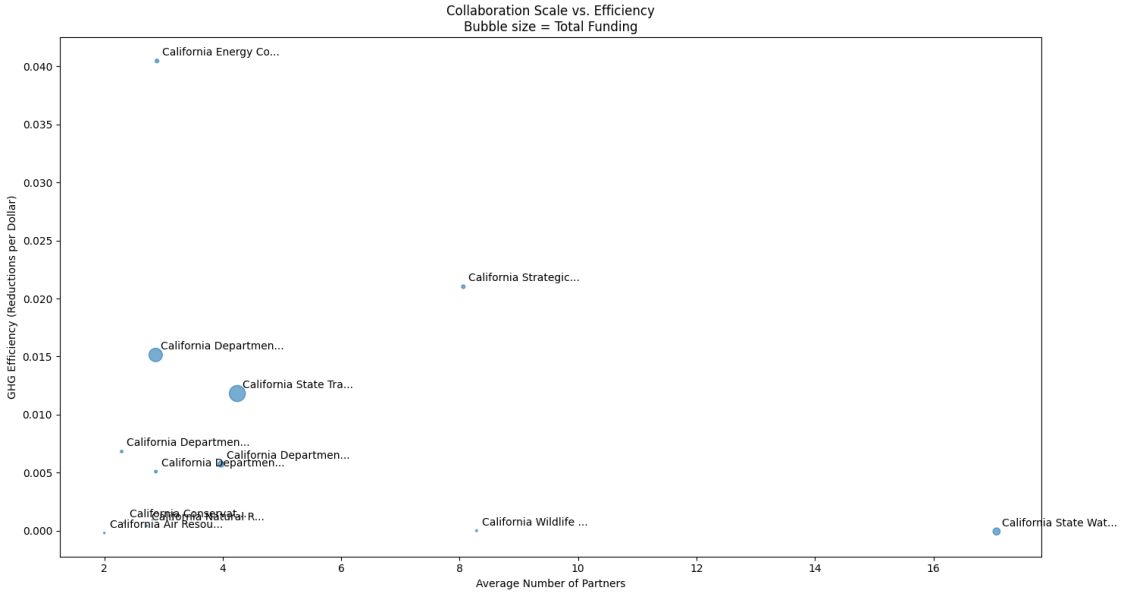
Agency Name	
California Air Resources Board	-413
California Conservation Corps	842
California Department of Food and Agriculture	42951
California Department of Forestry and Fire Prot…	185424
California Department of Resources Recycling an…	35434
California Department of Transportation	2460280
California Energy Commission	501007
California Natural Resources Agency	1191
California State Transportation Agency	2748646
California State Water Resources Control Board	-3234
California Strategic Growth Council	251149
California Wildlife Conservation Board	0

	Is Benefit Disadvantaged Communities \ mean
Agency Name	
California Air Resources Board	0.000
California Conservation Corps	0.000
California Department of Food and Agriculture	0.000
California Department of Forestry and Fire Prot…	0.168
California Department of Resources Recycling an…	0.130
California Department of Transportation	0.089
California Energy Commission	0.000
California Natural Resources Agency	0.000
California State Transportation Agency	0.105
California State Water Resources Control Board	0.000
California Strategic Growth Council	0.062
California Wildlife Conservation Board	0.000

	num_partners mean
Agency Name	
California Air Resources Board	2.000
California Conservation Corps	2.333

```
California Department of Food and Agriculture              2.293
California Department of Forestry and Fire Prot…         3.971
California Department of Resources Recycling an…         2.870
California Department of Transportation                    2.867
California Energy Commission                               2.889
California Natural Resources Agency                        2.714
California State Transportation Agency                     4.246
California State Water Resources Control Board           17.067
California Strategic Growth Council                        8.062
California Wildlife Conservation Board                     8.286
```

```
/tmp/ipykernel_164030/3430573634.py:6: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  multi_county['num_partners'] = multi_county['County'].str.count(',') + 1
```



[29]:
```
# 1. Characteristics of Successful Collaborations
# Define success metrics (GHG efficiency and DAC benefits)
multi_county['ghg_efficiency'] = multi_county['Total Project GHGReductions'] /␣
 ↪multi_county['Total Program GGRFFunding']

# Create success categories based on both metrics
def categorize_success(row):
    if row['ghg_efficiency'] > multi_county['ghg_efficiency'].median() and \
```

```
        row['Is Benefit Disadvantaged Communities'] > multi_county['Is Benefit␣
 ↪Disadvantaged Communities'].median():
            return 'High Performance'
        elif row['ghg_efficiency'] > multi_county['ghg_efficiency'].median():
            return 'High GHG Only'
        elif row['Is Benefit Disadvantaged Communities'] > multi_county['Is Benefit␣
 ↪Disadvantaged Communities'].median():
            return 'High Equity Only'
        else:
            return 'Below Median'

multi_county['performance_category'] = multi_county.apply(categorize_success,␣
 ↪axis=1)

# 2. Regional Analysis
# First, identify primary region for each collaboration
def get_primary_region(county_string):
    regions = {
        'Northern California': ['Shasta', 'Tehama', 'Butte', 'Plumas',␣
 ↪'Lassen', 'Modoc', 'Siskiyou', 'El Dorado', 'Nevada', 'Placer', 'Sierra'],
        'Central Coast': ['Monterey', 'San Benito', 'Santa Cruz', 'San Luis␣
 ↪Obispo'],
        'Sacramento Region': ['Sacramento', 'Yolo', 'Sutter', 'Yuba'],
        'Rural Interior': ['Inyo', 'Mono', 'Calaveras', 'Tuolumne'],
        'Multi-Region': []  # For projects with >5 counties
    }

    counties = county_string.split(',')
    if len(counties) > 5:
        return 'Multi-Region'

    # Rest of classification logic...

    region_counts = {region: 0 for region in regions.keys()}
    for county in counties:
        county = county.strip()
        for region, counties_list in regions.items():
            if county in counties_list:
                region_counts[region] += 1
                break


    return max(region_counts.items(), key=lambda x: x[1])[0]

multi_county['primary_region'] = multi_county['County'].
 ↪apply(get_primary_region)
```

```
# Analysis outputs
print("1. Success Characteristics Analysis:")
print("\nPerformance Category Distribution:")
print(multi_county['performance_category'].value_counts())

print("\nCharacteristics by Performance Category:")
success_chars = multi_county.groupby('performance_category').agg({
    'Total Program GGRFFunding': ['count', 'mean'],
    'num_partners': 'mean',
    'ghg_efficiency': 'mean',
    'Is Benefit Disadvantaged Communities': 'mean'
}).round(3)
print(success_chars)

print("\n2. Regional Collaboration Patterns:")
regional_patterns = multi_county.groupby('primary_region').agg({
    'Total Program GGRFFunding': ['count', 'sum', 'mean'],
    'Total Project GHGReductions': 'sum',
    'Is Benefit Disadvantaged Communities': 'mean',
    'num_partners': ['mean', 'max']
}).round(3)
print(regional_patterns)

# Visualize regional collaboration patterns
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(15, 12))

# 1. Project counts by region and performance
regional_performance = pd.crosstab(multi_county['primary_region'],
                                   multi_county['performance_category'])
regional_performance.plot(kind='bar', ax=ax1)
ax1.set_title('Performance Categories by Region')
ax1.set_ylabel('Number of Projects')

# 2. Average partners by region
multi_county.groupby('primary_region')['num_partners'].mean().plot(kind='bar',␣
 ↪ax=ax2)
ax2.set_title('Average Number of Partners by Region')
ax2.set_ylabel('Average Partners')

# 3. GHG efficiency by region
regional_ghg = multi_county.groupby('primary_region')['ghg_efficiency'].mean()
regional_ghg.plot(kind='bar', ax=ax3)
ax3.set_title('Average GHG Efficiency by Region')
ax3.set_ylabel('GHG Reductions per Dollar')

# 4. DAC benefit rate by region
```

```
regional_dac = multi_county.groupby('primary_region')['Is Benefit Disadvantaged␣
 ↪Communities'].mean()
regional_dac.plot(kind='bar', ax=ax4)
ax4.set_title('DAC Benefit Rate by Region')
ax4.set_ylabel('Proportion Benefiting DACs')

plt.tight_layout()
plt.show()
```

```
/tmp/ipykernel_164030/3413626539.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  multi_county['ghg_efficiency'] = multi_county['Total Project GHGReductions'] /
multi_county['Total Program GGRFFunding']
/tmp/ipykernel_164030/3413626539.py:17: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  multi_county['performance_category'] = multi_county.apply(categorize_success,
axis=1)
/tmp/ipykernel_164030/3413626539.py:47: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  multi_county['primary_region'] =
multi_county['County'].apply(get_primary_region)
1. Success Characteristics Analysis:

Performance Category Distribution:
performance_category
Below Median        475
High GHG Only       195
High Equity Only     49
High Performance     33
Name: count, dtype: int64

Characteristics by Performance Category:
                    Total Program GGRFFunding              num_partners  \
                                        count         mean         mean
performance_category
```

Below Median	475	308554.078	5.185
High Equity Only	49	60666.714	3.918
High GHG Only	195	1637050.159	3.154
High Performance	33	1585272.182	3.970

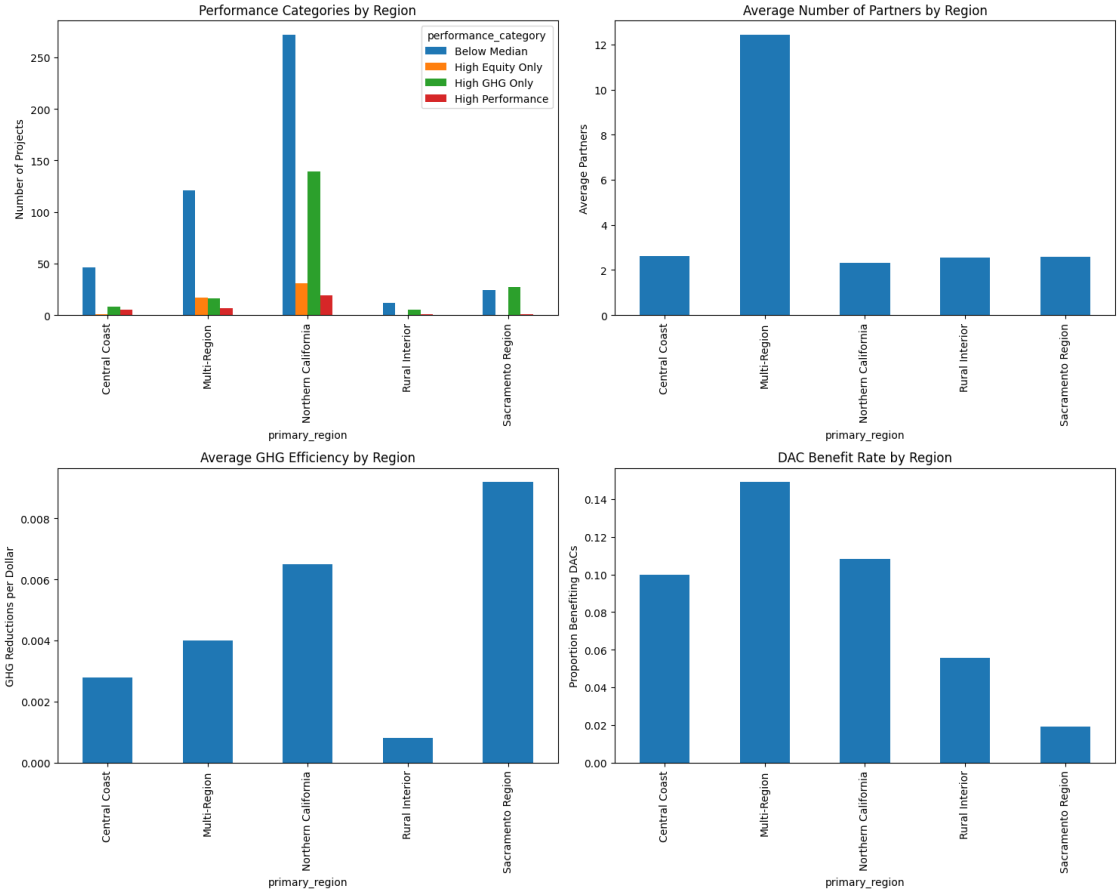
	ghg_efficiency	Is Benefit Disadvantaged Communities
	mean	mean
performance_category		
Below Median	0.000	0.0
High Equity Only	0.000	1.0
High GHG Only	0.012	0.0
High Performance	0.010	1.0

2. Regional Collaboration Patterns:

	Total Program GGRFFunding			\
	count	sum	mean	
primary_region				
Central Coast	60	19580751	326345.850	
Multi-Region	161	114266121	709727.460	
Northern California	461	370943021	804648.636	
Rural Interior	18	1364869	75826.056	
Sacramento Region	52	14919857	286920.327	

	Total Project GHGReductions	\
	sum	
primary_region		
Central Coast	312096	
Multi-Region	635116	
Northern California	5011802	
Rural Interior	745	
Sacramento Region	263518	

	Is Benefit Disadvantaged Communities	num_partners	
	mean	mean	max
primary_region			
Central Coast	0.100	2.600	4
Multi-Region	0.149	12.435	47
Northern California	0.108	2.304	5
Rural Interior	0.056	2.556	5
Sacramento Region	0.019	2.596	5



[30]:
```
# 1. Regional Success Models
# Look at top performing projects in each region
def analyze_regional_success():
    for region in multi_county['primary_region'].unique():
        region_data = multi_county[multi_county['primary_region'] == region]

        # Get top projects by different metrics
        top_ghg = region_data.nlargest(3, 'ghg_efficiency')
        top_equity = region_data.nlargest(3, 'Is Benefit Disadvantaged␣
 ↪Communities')

        print(f"\nTop Performers in {region}:")
        print("\nHighest GHG Efficiency Projects:")
        print(top_ghg[['Agency Name', 'Total Program GGRFFunding',
                       'ghg_efficiency', 'num_partners', 'County']].round(3))

        print("\nHighest DAC Benefit Projects:")
        print(top_equity[['Agency Name', 'Total Program GGRFFunding',
```

```
                        'Is Benefit Disadvantaged Communities',␣
 ↪'num_partners', 'County']].round(3))

# 2. Regional Characteristics Impact
def analyze_regional_factors():
    # Create regional metrics that might influence success
    regional_metrics = multi_county.groupby('primary_region').agg({
        'Total Program GGRFFunding': ['mean', 'count'],
        'num_partners': ['mean', 'max'],
        'ghg_efficiency': 'mean',
        'Is Benefit Disadvantaged Communities': 'mean'
    }).round(3)

    # Look at correlation between characteristics and outcomes
    print("\nRegional Characteristics Impact:")
    print(regional_metrics)

# 3. Agency Performance by Region
def analyze_agency_regional_performance():
    agency_region = multi_county.groupby(['Agency Name', 'primary_region']).
 ↪agg({
        'Total Program GGRFFunding': ['count', 'mean'],
        'ghg_efficiency': 'mean',
        'Is Benefit Disadvantaged Communities': 'mean',
        'num_partners': 'mean'
    }).round(3)

    print("\nAgency Performance by Region:")
    print(agency_region)

# 4. Temporal Analysis
def analyze_temporal_patterns():
    # Add year column if not already present
    multi_county['year'] = pd.to_datetime(multi_county['Date Operational']).dt.
 ↪year

    # Look at changes over time by region
    temporal = multi_county.groupby(['primary_region', 'year']).agg({
        'Total Program GGRFFunding': ['count', 'mean'],
        'ghg_efficiency': 'mean',
        'Is Benefit Disadvantaged Communities': 'mean',
        'num_partners': 'mean'
    }).round(3)

    print("\nTemporal Patterns by Region:")
    print(temporal)
```

```
# Visualize temporal trends
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(15, 12))

# Project counts over time by region
pivot_counts = pd.pivot_table(
    multi_county,
    values='Total Program GGRFFunding',
    index='year',
    columns='primary_region',
    aggfunc='count'
)
pivot_counts.plot(ax=ax1, marker='o')
ax1.set_title('Number of Projects Over Time by Region')

# Average project size over time
pivot_size = pd.pivot_table(
    multi_county,
    values='Total Program GGRFFunding',
    index='year',
    columns='primary_region',
    aggfunc='mean'
)
pivot_size.plot(ax=ax2, marker='o')
ax2.set_title('Average Project Size Over Time by Region')

# GHG efficiency over time
pivot_ghg = pd.pivot_table(
    multi_county,
    values='ghg_efficiency',
    index='year',
    columns='primary_region',
    aggfunc='mean'
)
pivot_ghg.plot(ax=ax3, marker='o')
ax3.set_title('GHG Efficiency Over Time by Region')

# DAC benefits over time
pivot_dac = pd.pivot_table(
    multi_county,
    values='Is Benefit Disadvantaged Communities',
    index='year',
    columns='primary_region',
    aggfunc='mean'
)
pivot_dac.plot(ax=ax4, marker='o')
ax4.set_title('DAC Benefit Rate Over Time by Region')
```

```
    plt.tight_layout()
    plt.show()

# Run all analyses
print("1. Regional Success Models:")
analyze_regional_success()
analyze_regional_factors()
analyze_agency_regional_performance()
analyze_temporal_patterns()
```

```
1. Regional Success Models:

Top Performers in Multi-Region:

Highest GHG Efficiency Projects:
                                    Agency Name  Total Program GGRFFunding  \
41296   California Department of Transportation                       20862
110839   California State Transportation Agency                      500000
100714   California State Transportation Agency                     6886000

        ghg_efficiency  num_partners  \
41296            0.036            14
110839           0.035             8
100714           0.025            12

                                                 County
41296   Amador, Butte, Colusa, El Dorado, Lake, Napa, …
110839  Alameda, Contra Costa, Napa, Placer, Sacrament…
100714  Alameda, Contra Costa, Fresno, Kern, Kings, Ma…

Highest DAC Benefit Projects:
                                          Agency Name  \
27603            California State Transportation Agency
41216  California Department of Forestry and Fire Pro…
41222  California Department of Forestry and Fire Pro…

       Total Program GGRFFunding  Is Benefit Disadvantaged Communities  \
27603                    1675000                                  True
41216                          0                                  True
41222                          0                                  True

       num_partners                                            County
27603             6  Los Angeles, Orange, San Diego, San Luis Obisp…
41216             8  Alameda, Contra Costa, Los Angeles, Merced, Sa…
41222             8  Los Angeles, Orange, Riverside, Sacramento, Sa…

Top Performers in Northern California:
```

```
Highest GHG Efficiency Projects:
                                    Agency Name  Total Program GGRFFunding  \
60417   California Department of Transportation                       645800
90888              California Energy Commission                       212629
100701   California State Transportation Agency                       666000

        ghg_efficiency  num_partners                              County
60417            0.114             3  Alameda, San Joaquin, Santa Clara
90888            0.111             2                Los Angeles, Solano
100701           0.111             2         San Francisco, Santa Clara

Highest DAC Benefit Projects:
                                            Agency Name  \
41220  California Department of Forestry and Fire Pro…
41231  California Department of Forestry and Fire Pro…
41233  California Department of Forestry and Fire Pro…

       Total Program GGRFFunding  Is Benefit Disadvantaged Communities  \
41220                          0                                   True
41231                          0                                   True
41233                          0                                   True

       num_partners                        County
41220             2          Imperial, San Diego
41231             2  Los Angeles, San Bernardino
41233             2    Riverside, San Bernardino

Top Performers in Central Coast:

Highest GHG Efficiency Projects:
                                             Agency Name  \
141413               California Strategic Growth Council
60272   California Department of Forestry and Fire Pro…
90896                       California Energy Commission

        Total Program GGRFFunding  ghg_efficiency  num_partners  \
141413                    4050000           0.062             2
60272                      192729           0.016             2
90896                     3440771           0.012             3

                                  County
141413                Monterey, San Benito
60272                Kern, San Luis Obispo
90896   Merced, San Luis Obispo, Stanislaus

Highest DAC Benefit Projects:
                                            Agency Name  \
41265  California Department of Resources Recycling a…
```

```
41316            California Department of Transportation
41317            California Department of Transportation

       Total Program GGRFFunding  Is Benefit Disadvantaged Communities  \
41265                    4000000                                   True
41316                      24983                                   True
41317                     110761                                   True

       num_partners                          County
41265             2    Los Angeles, San Luis Obispo
41316             2         San Benito, Santa Clara
41317             2  San Luis Obispo, Santa Barbara
```

Top Performers in Sacramento Region:

Highest GHG Efficiency Projects:

```
                                Agency Name  Total Program GGRFFunding  \
127475           California Energy Commission                     874375
127528  California State Transportation Agency                     907000
62284   California State Transportation Agency                    2666000

        ghg_efficiency  num_partners  \
127475           0.090             4
127528           0.029             5
62284            0.029             5

                                            County
127475        Los Angeles, Orange, Sacramento, San Diego
127528  Contra Costa, Sacramento, San Francisco, Solan…
62284   Contra Costa, Sacramento, San Francisco, Solan…
```

Highest DAC Benefit Projects:

```
                                          Agency Name  \
121037  California Department of Resources Recycling a…
41287            California Department of Transportation
60333            California Department of Transportation

        Total Program GGRFFunding  Is Benefit Disadvantaged Communities  \
121037                    1218026                                   True
41287                       10133                                  False
60333                       77847                                  False

        num_partners                    County
121037             2         Lake, Sacramento
41287              3     Colusa, Lake, Sutter
60333              3  Sacramento, Sutter, Yuba
```

Top Performers in Rural Interior:

Highest GHG Efficiency Projects:

	Agency Name	Total Program GGRFFunding	ghg_efficiency	num_partners	County
41330	California Department of Transportation	64588	0.004	5	Fresno, Inyo, Kern, Kings, Tulare
90188	California Department of Transportation	63613	0.004	3	Amador, Calaveras, Tuolumne
60413	California Department of Transportation	7949	0.003	2	Inyo, Mono

Highest DAC Benefit Projects:

	Agency Name	Total Program GGRFFunding	Is Benefit Disadvantaged Communities	num_partners	County
41330	California Department of Transportation	64588	True	5	Fresno, Inyo, Kern, Kings, Tulare
41361	California Department of Transportation	14390	False	2	Inyo, Mono
41363	California Department of Transportation	19938	False	2	Calaveras, Tuolumne

Regional Characteristics Impact:

	Total Program GGRFFunding		num_partners		ghg_efficiency	Is Benefit Disadvantaged Communities
	mean	count	mean	max	mean	mean
primary_region						
Central Coast	326345.850	60	2.600	4	0.003	0.100
Multi-Region	709727.460	161	12.435	47	0.004	0.149
Northern California	804648.636	461	2.304	5	0.006	0.108
Rural Interior	75826.056	18	2.556	5	0.001	0.056
Sacramento Region	286920.327	52	2.596	5	0.009	0.019

Agency Performance by Region:

Agency Name	primary_region	Total Program GGRFFunding \ count
California Air Resources Board	Northern California	25
California Conservation Corps	Central Coast	3
	Northern California	5
	Rural Interior	1
California Department of Food and Agriculture	Multi-Region	1
	Northern California	27
	Sacramento Region	13
California Department of Forestry and Fire Prot…	Central Coast	25
	Multi-Region	86
	Northern California	216
	Rural Interior	2
	Sacramento Region	10
California Department of Resources Recycling an…	Central Coast	1
	Northern California	21
	Sacramento Region	1
California Department of Transportation	Central Coast	16
	Multi-Region	10
	Northern California	100
	Rural Interior	15
	Sacramento Region	17
California Energy Commission	Central Coast	1
	Northern California	7

```
                                                    Sacramento Region
1
California Natural Resources Agency                 Central Coast
5
                                                    Northern California
2
California State Transportation Agency              Multi-Region
11
                                                    Northern California
42
                                                    Sacramento Region
4
California State Water Resources Control Board      Central Coast
7
                                                    Multi-Region
35
                                                    Northern California
3
California Strategic Growth Council                 Central Coast
1
                                                    Multi-Region
14
                                                    Northern California
11
                                                    Sacramento Region
6
California Wildlife Conservation Board              Central Coast
1
                                                    Multi-Region
4
                                                    Northern California
2

     \
mean
Agency Name                                         primary_region
California Air Resources Board                      Northern California
84161.920
California Conservation Corps                       Central Coast
176120.667
                                                    Northern California
112423.600
                                                    Rural Interior
157873.000
California Department of Food and Agriculture       Multi-Region
1967590.000
                                                    Northern California
105924.259
```

	Sacramento Region	113025.077
California Department of Forestry and Fire Prot…	Central Coast	35336.480
	Multi-Region	102608.407
	Northern California	99839.486
	Rural Interior	351529.000
	Sacramento Region	37538.600
California Department of Resources Recycling an…	Central Coast	4000000.000
	Northern California	83056.095
	Sacramento Region	1218026.000
California Department of Transportation	Central Coast	124621.812
	Multi-Region	1221953.800
	Northern California	1421644.780
	Rural Interior	33595.867
	Sacramento Region	330789.765
California Energy Commission	Central Coast	3440771.000
	Northern California	1152074.286
	Sacramento Region	874375.000
California Natural Resources Agency	Central Coast	396888.600
	Northern California	447352.500
California State Transportation Agency	Multi-Region	3637818.182
	Northern California	4476261.905
	Sacramento Region	1142750.000
California State Water Resources Control Board	Central Coast	99973.429
	Multi-Region	1285884.486

```
                                                    Northern California
186410.000
California Strategic Growth Council                 Central Coast
4050000.000
                                                    Multi-Region
379136.643
                                                    Northern California
163823.636
                                                    Sacramento Region
131386.333
California Wildlife Conservation Board              Central Coast
2000000.000
                                                    Multi-Region
231200.000
                                                    Northern California
309700.000

ghg_efficiency  \
 mean
Agency Name                                         primary_region
California Air Resources Board                      Northern California
0.000
California Conservation Corps                       Central Coast
0.000
                                                    Northern California
0.001
                                                    Rural Interior
0.001
California Department of Food and Agriculture       Multi-Region
0.000
                                                    Northern California
0.011
                                                    Sacramento Region
0.012
California Department of Forestry and Fire Prot… Central Coast
0.006
                                                    Multi-Region
0.006
                                                    Northern California
0.004
                                                    Rural Interior
0.000
                                                    Sacramento Region
0.003
California Department of Resources Recycling an… Central Coast
0.004
                                                    Northern California
0.003
```

		Sacramento Region
0.009		
California Department of Transportation		Central Coast
0.001		
		Multi-Region
0.007		
		Northern California
0.005		
		Rural Interior
0.001		
		Sacramento Region
0.002		
California Energy Commission		Central Coast
0.012		
		Northern California
0.044		
		Sacramento Region
0.090		
California Natural Resources Agency		Central Coast
0.000		
		Northern California
0.001		
California State Transportation Agency		Multi-Region
0.019		
		Northern California
0.012		
		Sacramento Region
0.024		
California State Water Resources Control Board		Central Coast
-0.001		
		Multi-Region
-0.000		
		Northern California
-0.000		
California Strategic Growth Council		Central Coast
0.062		
		Multi-Region
0.000		
		Northern California
0.000		
		Sacramento Region
0.000		
California Wildlife Conservation Board		Central Coast
0.000		
		Multi-Region
0.000		
		Northern California
0.000		

```
                                                                    Is
Benefit Disadvantaged Communities  \
                        mean
Agency Name                                          primary_region
California Air Resources Board                       Northern California
0.000
California Conservation Corps                        Central Coast
0.000
                                                     Northern California
0.000
                                                     Rural Interior
0.000
California Department of Food and Agriculture        Multi-Region
0.000
                                                     Northern California
0.000
                                                     Sacramento Region
0.000
California Department of Forestry and Fire Prot… Central Coast
0.120
                                                     Multi-Region
0.198
                                                     Northern California
0.171
                                                     Rural Interior
0.000
                                                     Sacramento Region
0.000
California Department of Resources Recycling an… Central Coast
1.000
                                                     Northern California
0.048
                                                     Sacramento Region
1.000
California Department of Transportation              Central Coast
0.125
                                                     Multi-Region
0.600
                                                     Northern California
0.050
                                                     Rural Interior
0.067
                                                     Sacramento Region
0.000
California Energy Commission                         Central Coast
0.000
                                                     Northern California
```

```
0.000
                                                    Sacramento Region
0.000
California Natural Resources Agency                 Central Coast
0.000
                                                    Northern California
0.000
California State Transportation Agency              Multi-Region
0.091
                                                    Northern California
0.119
                                                    Sacramento Region
0.000
California State Water Resources Control Board      Central Coast
0.000
                                                    Multi-Region
0.000
                                                    Northern California
0.000
California Strategic Growth Council                 Central Coast
0.000
                                                    Multi-Region
0.000
                                                    Northern California
0.182
                                                    Sacramento Region
0.000
California Wildlife Conservation Board              Central Coast
0.000
                                                    Multi-Region
0.000
                                                    Northern California
0.000

num_partners
mean
Agency Name                                         primary_region
California Air Resources Board                      Northern California
2.000
California Conservation Corps                       Central Coast
2.000
                                                    Northern California
2.400
                                                    Rural Interior
3.000
California Department of Food and Agriculture       Multi-Region
10.000
                                                    Northern California
```

2.074
Sacramento Region
2.154
California Department of Forestry and Fire Prot… Central Coast
2.320
Multi-Region
9.407
Northern California
2.097
Rural Interior
3.000
Sacramento Region
2.000
California Department of Resources Recycling an… Central Coast
2.000
Northern California
2.952
Sacramento Region
2.000
California Department of Transportation Central Coast
2.500
Multi-Region
9.500
Northern California
2.420
Rural Interior
2.467
Sacramento Region
2.294
California Energy Commission Central Coast
3.000
Northern California
2.714
Sacramento Region
4.000
California Natural Resources Agency Central Coast
3.000
Northern California
2.000
California State Transportation Agency Multi-Region
10.636
Northern California
2.500
Sacramento Region
5.000
California State Water Resources Control Board Central Coast
4.000
Multi-Region

```
20.886
                                                    Northern California
3.000
California Strategic Growth Council                 Central Coast
2.000
                                                    Multi-Region
13.714
                                                    Northern California
3.818
                                                    Sacramento Region
3.667
California Wildlife Conservation Board              Central Coast
2.000
                                                    Multi-Region
12.000
                                                    Northern California
4.000

Temporal Patterns by Region:
                        Total Program GGRFFunding                 \
                                              count          mean
primary_region      year
Central Coast       2017                          5   165285.400
                    2018                          2   138319.000
                    2019                          8   641301.875
                    2020                         32    77539.219
                    2021                          7   121835.714
                    2022                          3  1469704.333
                    2023                          2  1081641.000
                    2024                          1  3440771.000
Multi-Region        2015                         16        0.000
                    2016                          1  1675000.000
                    2017                         12    62775.417
                    2018                          5   997038.400
                    2019                         11   431204.000
                    2020                         59  1042264.475
                    2021                         12   420457.833
                    2022                         25  1170613.680
                    2023                          7   292694.429
                    2024                         13   327390.692
Northern California 2015                         16        0.000
                    2016                          2   142220.500
                    2017                         36   101603.861
                    2018                         23   715546.043
                    2019                         47  1041667.234
                    2020                         60   874110.533
                    2021                         97   659413.691
                    2022                         99  1179907.667
```

	2023	62	807812.710
	2024	19	962100.789
Rural Interior	2017	2	17164.000
	2018	3	34777.333
	2019	1	42431.000
	2020	3	36044.333
	2021	2	27071.000
	2022	2	40022.000
	2023	5	188291.800
Sacramento Region	2017	5	124026.600
	2018	4	156487.750
	2019	7	850873.286
	2020	4	222162.500
	2021	22	74069.045
	2022	2	233685.500
	2023	6	559790.833
	2024	2	686687.500

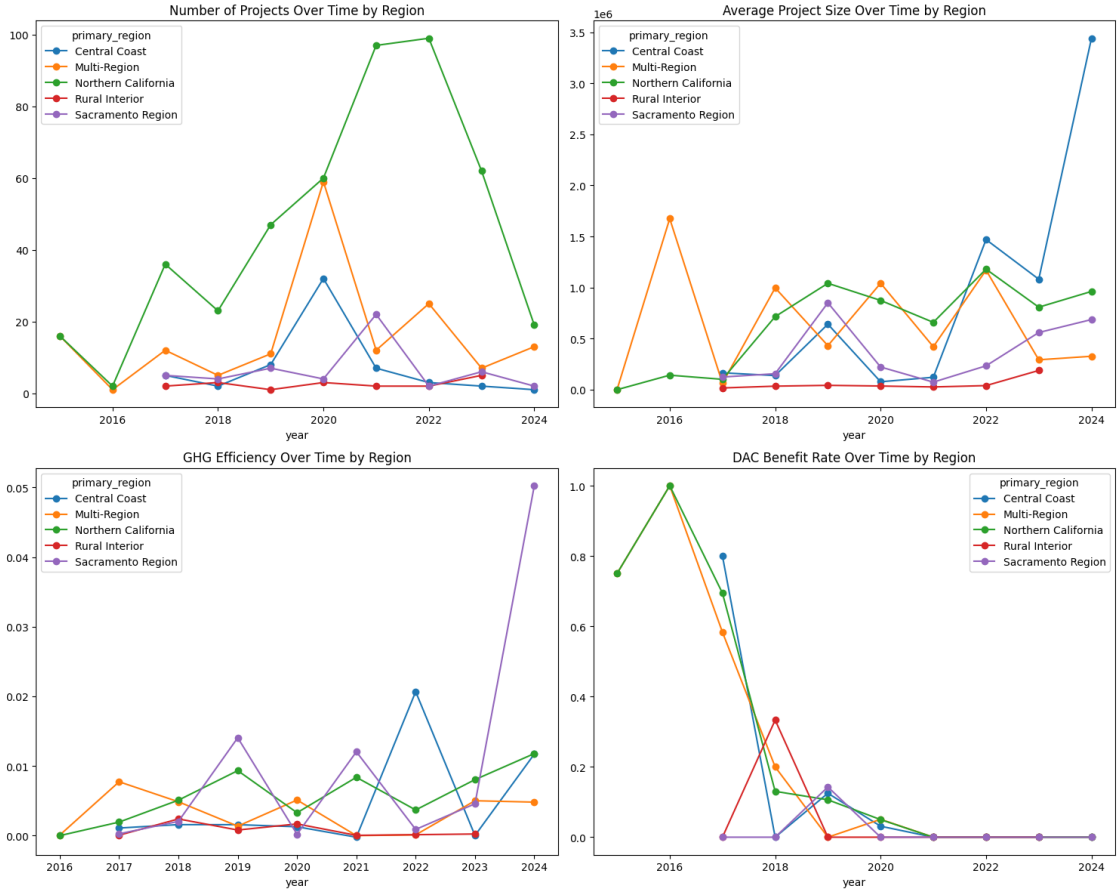
		ghg_efficiency Is Benefit	Disadvantaged Communities \
		mean	mean
primary_region	year		
Central Coast	2017	0.001	0.800
	2018	0.002	0.000
	2019	0.002	0.125
	2020	0.001	0.031
	2021	-0.000	0.000
	2022	0.021	0.000
	2023	0.000	0.000
	2024	0.012	0.000
Multi-Region	2015	NaN	0.750
	2016	0.000	1.000
	2017	0.008	0.583
	2018	0.005	0.200
	2019	0.001	0.000
	2020	0.005	0.051
	2021	0.000	0.000
	2022	0.000	0.000
	2023	0.005	0.000
	2024	0.005	0.000
Northern California	2015	NaN	0.750
	2016	0.000	1.000
	2017	0.002	0.694
	2018	0.005	0.130
	2019	0.009	0.106
	2020	0.003	0.050
	2021	0.008	0.000
	2022	0.004	0.000
	2023	0.008	0.000

	2024	0.012	0.000
Rural Interior	2017	0.000	0.000
	2018	0.002	0.333
	2019	0.001	0.000
	2020	0.002	0.000
	2021	0.000	0.000
	2022	0.000	0.000
	2023	0.000	0.000
Sacramento Region	2017	0.000	0.000
	2018	0.002	0.000
	2019	0.014	0.143
	2020	0.000	0.000
	2021	0.012	0.000
	2022	0.001	0.000
	2023	0.005	0.000
	2024	0.050	0.000

		num_partners
		mean
primary_region	year	
Central Coast	2017	2.200
	2018	2.000
	2019	3.000
	2020	2.562
	2021	3.143
	2022	2.000
	2023	2.000
	2024	3.000
Multi-Region	2015	7.500
	2016	6.000
	2017	9.167
	2018	9.800
	2019	8.000
	2020	16.407
	2021	16.083
	2022	10.240
	2023	9.286
	2024	11.308
Northern California	2015	2.000
	2016	4.000
	2017	2.333
	2018	2.304
	2019	2.319
	2020	2.067
	2021	2.144
	2022	2.404
	2023	2.677
	2024	2.105

```
Rural Interior      2017        2.000
                    2018        3.000
                    2019        3.000
                    2020        2.333
                    2021        3.000
                    2022        2.000
                    2023        2.600
Sacramento Region   2017        3.800
                    2018        2.750
                    2019        3.571
                    2020        2.000
                    2021        2.045
                    2022        2.000
                    2023        2.333
                    2024        4.500
```

```
/tmp/ipykernel_164030/2787321163.py:49: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  multi_county['year'] = pd.to_datetime(multi_county['Date
Operational']).dt.year
```



[31]:
```python
# Identify programs that maintained performance through transitions
def analyze_transitions():
    # Add period categorization
    multi_county['period'] = pd.cut(
        multi_county['year'],
        bins=[2014, 2017, 2020, 2024],
        labels=['Early', 'Growth', 'Recent']
    )

    # Look at programs that maintained both GHG and DAC benefits
    transitions = multi_county.groupby(['Agency Name', 'period']).agg({
        'Total Program GGRFFunding': ['count', 'mean'],
        'ghg_efficiency': 'mean',
        'Is Benefit Disadvantaged Communities': 'mean',
        'num_partners': 'mean'
    }).round(3)

    print("Programs Through Transitions:")
    print(transitions)
```

```
    # Identify consistently high performers
    high_performers = transitions[
        (transitions[('ghg_efficiency', 'mean')] >␣
 ↪transitions[('ghg_efficiency', 'mean')].median()) &
        (transitions[('Is Benefit Disadvantaged Communities', 'mean')] > 0.1)
    ]

    print("\nConsistently High Performing Programs:")
    print(high_performers)

# 2. 2020 Analysis
def analyze_2020_pivot():
    pre_2020 = multi_county[multi_county['year'] < 2020]
    post_2020 = multi_county[multi_county['year'] >= 2020]

    metrics = ['Total Program GGRFFunding', 'ghg_efficiency',
               'Is Benefit Disadvantaged Communities', 'num_partners']

    comparison = pd.DataFrame({
        'Pre-2020': pre_2020[metrics].mean(),
        'Post-2020': post_2020[metrics].mean(),
        'Change%': ((post_2020[metrics].mean() - pre_2020[metrics].mean()) /
                    pre_2020[metrics].mean() * 100)
    }).round(2)

    print("\n2020 Pivot Analysis:")
    print(comparison)

# 3. Regional Adaptation
def analyze_regional_adaptation():
    # Look at how regions adapted their collaboration approaches
    regional_evolution = multi_county.groupby(['primary_region', 'year']).agg({
        'num_partners': ['mean', 'max'],
        'Total Program GGRFFunding': ['count', 'mean'],
        'ghg_efficiency': 'mean',
        'Is Benefit Disadvantaged Communities': 'mean'
    }).round(3)

    # Plot evolution of collaboration strategies
    fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(15, 12))

    # Partner network evolution
    partner_evolution = pd.pivot_table(
        multi_county,
        values='num_partners',
        index='year',
```

```
        columns='primary_region',
        aggfunc='mean'
    )
    partner_evolution.plot(ax=ax1, marker='o')
    ax1.set_title('Evolution of Partnership Size by Region')

    # Project scale evolution
    scale_evolution = pd.pivot_table(
        multi_county,
        values='Total Program GGRFFunding',
        index='year',
        columns='primary_region',
        aggfunc='mean'
    )/1e6
    scale_evolution.plot(ax=ax2, marker='o')
    ax2.set_title('Evolution of Project Scale by Region (Millions $)')

    # Success metrics evolution
    success_evolution = pd.pivot_table(
        multi_county,
        values=['ghg_efficiency', 'Is Benefit Disadvantaged Communities'],
        index='year',
        columns='primary_region',
        aggfunc='mean'
    )

    success_evolution['ghg_efficiency'].plot(ax=ax3, marker='o')
    ax3.set_title('Evolution of GHG Efficiency')

    success_evolution['Is Benefit Disadvantaged Communities'].plot(ax=ax4,␣
 ↪marker='o')
    ax4.set_title('Evolution of DAC Benefits')

    plt.tight_layout()
    plt.show()

# 4. Agency Transitions
def analyze_agency_transitions():
    agency_evolution = multi_county.groupby(['Agency Name', 'year']).agg({
        'Total Program GGRFFunding': ['count', 'mean'],
        'ghg_efficiency': 'mean',
        'Is Benefit Disadvantaged Communities': 'mean',
        'num_partners': 'mean'
    }).round(3)

    print("\nAgency Evolution Patterns:")
    print(agency_evolution)
```

```
# Run all analyses
analyze_transitions()
analyze_2020_pivot()
analyze_regional_adaptation()
analyze_agency_transitions()
```

```
/tmp/ipykernel_164030/883521891.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  multi_county['period'] = pd.cut(
/tmp/ipykernel_164030/883521891.py:11: FutureWarning: The default of
observed=False is deprecated and will be changed to True in a future version of
pandas. Pass observed=False to retain current behavior or observed=True to adopt
the future default and silence this warning.
  transitions = multi_county.groupby(['Agency Name', 'period']).agg({
```

```
Programs Through Transitions:
                                                        Total Program
GGRFFunding  \
count
Agency Name                                        period
California Air Resources Board                     Early
0
                                                   Growth
13
                                                   Recent
12
California Conservation Corps                      Early
0
                                                   Growth
2
                                                   Recent
7
California Department of Food and Agriculture      Early
0
                                                   Growth
1
                                                   Recent
40
California Department of Forestry and Fire Prot… Early
65
                                                   Growth
105
                                                   Recent
```

169
California Department of Resources Recycling an… Early
0
Growth
3
Recent
20
California Department of Transportation Early
17
Growth
79
Recent
62
California Energy Commission Early
0
Growth
0
Recent
9
California Natural Resources Agency Early
0
Growth
7
Recent
0
California State Transportation Agency Early
1
Growth
26
Recent
30
California State Water Resources Control Board Early
0
Growth
18
Recent
27
California Strategic Growth Council Early
12
Growth
15
Recent
5
California Wildlife Conservation Board Early
0
Growth
0
Recent

7

Agency Name	period	mean
California Air Resources Board	Early	NaN
	Growth	122263.000
	Recent	42885.750
California Conservation Corps	Early	NaN
	Growth	105587.500
	Recent	148168.286
California Department of Food and Agriculture	Early	NaN
	Growth	1967590.000
	Recent	108232.025
California Department of Forestry and Fire Prot…	Early	53680.262
	Growth	191617.610
	Recent	51730.426
California Department of Resources Recycling an…	Early	NaN
	Growth	1919908.667
	Recent	60123.900
California Department of Transportation	Early	52512.647
	Growth	736585.089
	Recent	1668103.097
California Energy Commission	Early	NaN
	Growth	NaN
	Recent	1375518.444
California Natural Resources Agency	Early	NaN
	Growth	411306.857
	Recent	NaN
California State Transportation Agency	Early	1675000.000
	Growth	3323846.154
	Recent	4816500.000
California State Water Resources Control Board	Early	NaN
	Growth	1225722.556
	Recent	896370.185
California Strategic Growth Council	Early	149536.750
	Growth	366558.267
	Recent	931095.200
California Wildlife Conservation Board	Early	NaN
	Growth	NaN
	Recent	506314.286

Agency Name	period	ghg_efficiency mean
California Air Resources Board	Early	NaN
	Growth	0.000
	Recent	-0.000

California Conservation Corps	Early	NaN
	Growth	0.000
	Recent	0.001
California Department of Food and Agriculture	Early	NaN
	Growth	0.000
	Recent	0.011
California Department of Forestry and Fire Prot…	Early	0.003
	Growth	0.007
	Recent	0.003
California Department of Resources Recycling an…	Early	NaN
	Growth	0.010
	Recent	0.003
California Department of Transportation	Early	0.004
	Growth	0.005
	Recent	0.003
California Energy Commission	Early	NaN
	Growth	NaN
	Recent	0.046
California Natural Resources Agency	Early	NaN
	Growth	0.000
	Recent	NaN
California State Transportation Agency	Early	0.000
	Growth	0.017
	Recent	0.012
California State Water Resources Control Board	Early	NaN
	Growth	-0.000
	Recent	-0.000
California Strategic Growth Council	Early	0.000
	Growth	0.000
	Recent	0.012
California Wildlife Conservation Board	Early	NaN
	Growth	NaN
	Recent	0.000

		Is Benefit Disadvantaged Communities \ mean
Agency Name	period	
California Air Resources Board	Early	NaN
	Growth	0.000
	Recent	0.000
California Conservation Corps	Early	NaN
	Growth	0.000

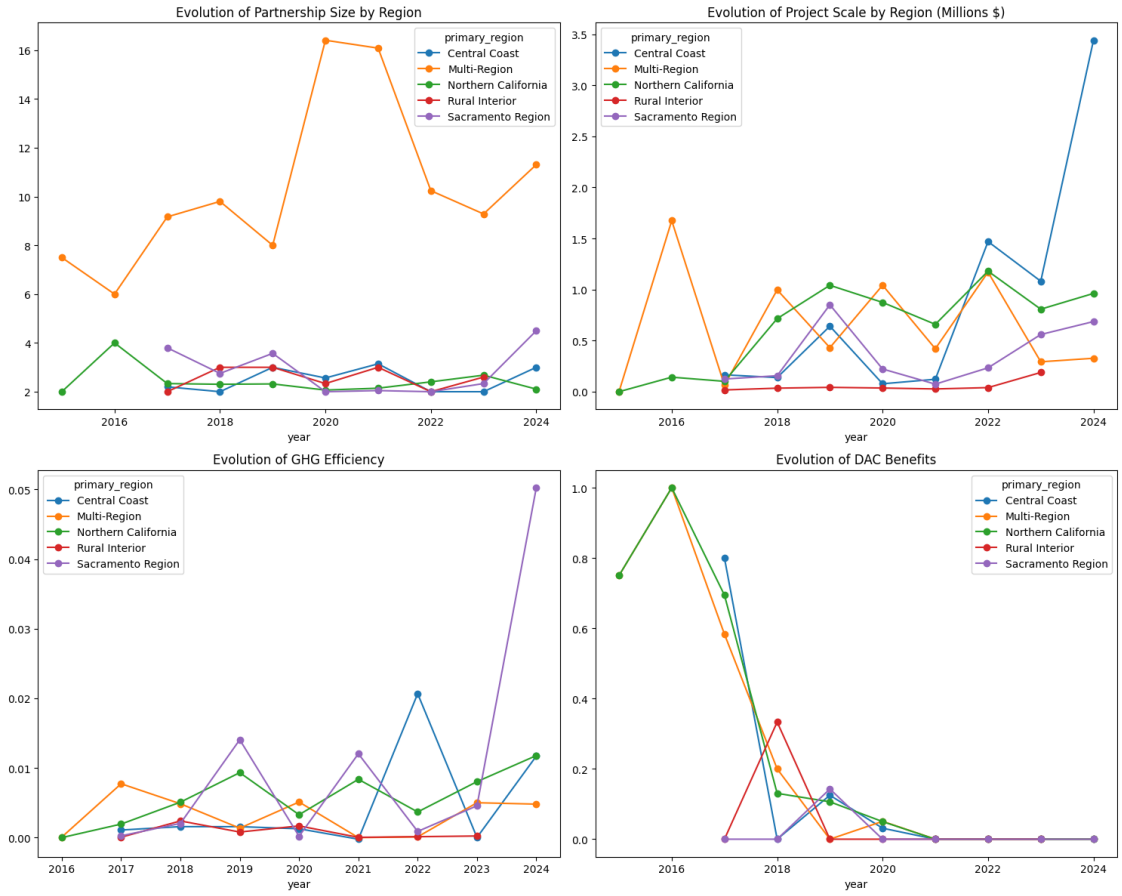
```
                                                    Recent
0.000
California Department of Food and Agriculture      Early
NaN
                                                    Growth
0.000
                                                    Recent
0.000
California Department of Forestry and Fire Prot… Early
0.769
                                                    Growth
0.067
                                                    Recent
0.000
California Department of Resources Recycling an… Early
NaN
                                                    Growth
1.000
                                                    Recent
0.000
California Department of Transportation            Early
0.588
                                                    Growth
0.051
                                                    Recent
0.000
California Energy Commission                       Early
NaN
                                                    Growth
NaN
                                                    Recent
0.000
California Natural Resources Agency                Early
NaN
                                                    Growth
0.000
                                                    Recent
NaN
California State Transportation Agency             Early
1.000
                                                    Growth
0.192
                                                    Recent
0.000
California State Water Resources Control Board     Early
NaN
                                                    Growth
0.000
```

	Recent	0.000
California Strategic Growth Council	Early	0.167
	Growth	0.000
	Recent	0.000
California Wildlife Conservation Board	Early	NaN
	Growth	NaN
	Recent	0.000

		num_partners
		mean
Agency Name	period	
California Air Resources Board	Early	NaN
	Growth	2.000
	Recent	2.000
California Conservation Corps	Early	NaN
	Growth	2.000
	Recent	2.429
California Department of Food and Agriculture	Early	NaN
	Growth	10.000
	Recent	2.100
California Department of Forestry and Fire Prot…	Early	3.769
	Growth	4.410
	Recent	3.775
California Department of Resources Recycling an…	Early	NaN
	Growth	2.000
	Recent	3.000
California Department of Transportation	Early	4.882
	Growth	2.835
	Recent	2.355
California Energy Commission	Early	NaN
	Growth	NaN
	Recent	2.889
California Natural Resources Agency	Early	NaN
	Growth	2.714
	Recent	NaN
California State Transportation Agency	Early	6.000
	Growth	5.462
	Recent	3.133
California State Water Resources Control Board	Early	NaN
	Growth	27.222
	Recent	10.296

```
California Strategic Growth Council                Early         5.000
                                                   Growth       12.000
                                                   Recent        3.600
California Wildlife Conservation Board             Early           NaN
                                                   Growth          NaN
                                                   Recent        8.286

Consistently High Performing Programs:
                                                                 Total Program
GGRFFunding  \
count
Agency Name                                        period
California Department of Resources Recycling an… Growth
3
California Department of Transportation            Early
17
California State Transportation Agency             Growth
26

                                                                              \
                                                                       mean
Agency Name                                        period
California Department of Resources Recycling an… Growth  1919908.667
California Department of Transportation            Early     52512.647
California State Transportation Agency             Growth  3323846.154

                                                              ghg_efficiency  \
                                                                        mean
Agency Name                                        period
California Department of Resources Recycling an… Growth          0.010
California Department of Transportation            Early           0.004
California State Transportation Agency             Growth          0.017

                                                              Is Benefit
Disadvantaged Communities  \
          mean
Agency Name                                        period
California Department of Resources Recycling an… Growth
1.000
California Department of Transportation            Early
0.588
California State Transportation Agency             Growth
0.192

                                                              num_partners
                                                                      mean
Agency Name                                        period
California Department of Resources Recycling an… Growth        2.000
```

```
California Department of Transportation              Early         4.882
California State Transportation Agency               Growth        5.462

2020 Pivot Analysis:
                                          Pre-2020   Post-2020   Change%
Total Program GGRFFunding                461803.92   780115.40     68.93
ghg_efficiency                                0.01        0.01      6.24
Is Benefit Disadvantaged Communities          0.36        0.01    -96.48
num_partners                                  3.73        4.82     29.13
```




```
Agency Evolution Patterns:
                                                        Total Program
GGRFFunding  \
count
Agency Name                                          year
California Air Resources Board                       2018
3
                                                     2019
3
```

	2020	
7		
	2021	
3		
	2022	
9		
California Conservation Corps	2019	
2		
	2021	
2		
	2022	
3		
	2023	
2		
California Department of Food and Agriculture	2020	
1		
	2021	
25		
	2023	
15		
California Department of Forestry and Fire Prot…	2015	
32		
	2017	
33		
	2018	
9		
	2019	
17		
	2020	
79		
	2021	
68		
	2022	
75		
	2023	
7		
	2024	
19		
California Department of Resources Recycling an…	2018	
1		
	2019	
2		
	2021	
2		
	2023	
18		
California Department of Transportation	2017	
17		

	2018	
17		
	2019	
33		
	2020	
29		
	2021	
17		
	2022	
21		
	2023	
23		
	2024	
1		
California Energy Commission	2022	
1		
	2023	
3		
	2024	
5		
California Natural Resources Agency	2020	
7		
California State Transportation Agency	2016	
1		
	2018	
2		
	2019	
14		
	2020	
10		
	2021	
5		
	2022	
10		
	2023	
6		
	2024	
9		
California State Water Resources Control Board	2020	
18		
	2021	
16		
	2022	
8		
	2023	
2		
	2024	
1		

```
California Strategic Growth Council                2016
2
                                                   2017
10
                                                   2018
5
                                                   2019
3
                                                   2020
7
                                                   2021
2
                                                   2022
3
California Wildlife Conservation Board             2022
1
                                                   2023
6

                                                                \
                                                          mean
Agency Name                                        year
California Air Resources Board                     2018   365621.000
                                                   2019    49176.000
                                                   2020    49289.714
                                                   2021    34991.667
                                                   2022    45517.111
California Conservation Corps                      2019   105587.500
                                                   2021   143363.500
                                                   2022   141773.333
                                                   2023   162565.500
California Department of Food and Agriculture      2020  1967590.000
                                                   2021   125553.720
                                                   2023    79362.533
California Department of Forestry and Fire Prot… 2015        0.000
                                                   2017   105733.848
                                                   2018   621380.222
                                                   2019    17530.647
                                                   2020   180119.063
                                                   2021    14701.985
                                                   2022    19334.160
                                                   2023   241527.429
                                                   2024   242208.053
California Department of Resources Recycling an… 2018   541700.000
                                                   2019  2609013.000
                                                   2021   494317.500
                                                   2023    11880.167
California Department of Transportation            2017    52512.647
```

	2018	252836.824
	2019	1024964.030
	2020	692006.310
	2021	984323.294
	2022	2619571.714
	2023	1372278.261
	2024	115490.000
California Energy Commission	2022	459624.000
	2023	1648954.000
	2024	1394636.000
California Natural Resources Agency	2020	411306.857
California State Transportation Agency	2016	1675000.000
	2018	4465000.000
	2019	1666571.429
	2020	5415800.000
	2021	8639600.000
	2022	7232800.000
	2023	2427666.667
	2024	1600333.333
California State Water Resources Control Board	2020	1225722.556
	2021	347358.562
	2022	2082571.500
	2023	363584.500
	2024	1256517.000
California Strategic Growth Council	2016	142220.500
	2017	151000.000
	2018	398092.200
	2019	600000.000
	2020	243987.571
	2021	268492.500
	2022	1372830.333
California Wildlife Conservation Board	2022	170000.000
	2023	562366.667

		ghg_efficiency \
		mean
Agency Name	year	
California Air Resources Board	2018	0.000
	2019	0.001
	2020	0.000
	2021	-0.003
	2022	0.001
California Conservation Corps	2019	0.000
	2021	0.000
	2022	0.001
	2023	0.002
California Department of Food and Agriculture	2020	0.000
	2021	0.011

	2023	0.011
California Department of Forestry and Fire Prot…	2015	NaN
	2017	0.003
	2018	0.007
	2019	0.005
	2020	0.006
	2021	0.015
	2022	0.003
	2023	0.000
	2024	0.005
California Department of Resources Recycling an…	2018	0.016
	2019	0.006
	2021	0.001
	2023	0.003
California Department of Transportation	2017	0.004
	2018	0.004
	2019	0.008
	2020	0.001
	2021	0.003
	2022	0.004
	2023	0.002
	2024	0.002
California Energy Commission	2022	0.004
	2023	0.076
	2024	0.036
California Natural Resources Agency	2020	0.000
California State Transportation Agency	2016	0.000
	2018	0.007
	2019	0.015
	2020	0.022
	2021	0.029
	2022	0.007
	2023	0.011
	2024	0.010
California State Water Resources Control Board	2020	-0.000
	2021	-0.000
	2022	-0.000
	2023	0.000
	2024	-0.000
California Strategic Growth Council	2016	0.000
	2017	0.000
	2018	0.000
	2019	0.000
	2020	0.000
	2021	0.000
	2022	0.021
California Wildlife Conservation Board	2022	0.000
	2023	0.000

```
                                                        Is Benefit Disadvantaged
Communities  \
        mean
Agency Name                                          year
California Air Resources Board                       2018
0.000
                                                     2019
0.000
                                                     2020
0.000
                                                     2021
0.000
                                                     2022
0.000
California Conservation Corps                        2019
0.000
                                                     2021
0.000
                                                     2022
0.000
                                                     2023
0.000
California Department of Food and Agriculture       2020
0.000
                                                     2021
0.000
                                                     2023
0.000
California Department of Forestry and Fire Prot… 2015
0.750
                                                     2017
0.788
                                                     2018
0.000
                                                     2019
0.059
                                                     2020
0.076
                                                     2021
0.000
                                                     2022
0.000
                                                     2023
0.000
                                                     2024
0.000
California Department of Resources Recycling an… 2018
```

```
1.000
                                                  2019
1.000
                                                  2021
0.000
                                                  2023
0.000
California Department of Transportation           2017
0.588
                                                  2018
0.118
                                                  2019
0.061
                                                  2020
0.000
                                                  2021
0.000
                                                  2022
0.000
                                                  2023
0.000
                                                  2024
0.000
California Energy Commission                      2022
0.000
                                                  2023
0.000
                                                  2024
0.000
California Natural Resources Agency               2020
0.000
California State Transportation Agency            2016
1.000
                                                  2018
1.000
                                                  2019
0.143
                                                  2020
0.100
                                                  2021
0.000
                                                  2022
0.000
                                                  2023
0.000
                                                  2024
0.000
California State Water Resources Control Board    2020
```

```
0.000
                                                2021
0.000
                                                2022
0.000
                                                2023
0.000
                                                2024
0.000
California Strategic Growth Council             2016
1.000
                                                2017
0.000
                                                2018
0.000
                                                2019
0.000
                                                2020
0.000
                                                2021
0.000
                                                2022
0.000
California Wildlife Conservation Board          2022
0.000
                                                2023
0.000
```

		num_partners
		mean
Agency Name	year	
California Air Resources Board	2018	2.000
	2019	2.000
	2020	2.000
	2021	2.000
	2022	2.000
California Conservation Corps	2019	2.000
	2021	2.000
	2022	2.333
	2023	3.000
California Department of Food and Agriculture	2020	10.000
	2021	2.040
	2023	2.200
California Department of Forestry and Fire Prot…	2015	4.750
	2017	2.818
	2018	3.889
	2019	4.588
	2020	4.430

	2021	2.000
	2022	4.280
	2023	4.143
	2024	8.000
California Department of Resources Recycling an…	2018	2.000
	2019	2.000
	2021	3.000
	2023	3.000
California Department of Transportation	2017	4.882
	2018	3.059
	2019	2.667
	2020	2.897
	2021	2.471
	2022	2.190
	2023	2.435
	2024	2.000
California Energy Commission	2022	3.000
	2023	2.667
	2024	3.000
California Natural Resources Agency	2020	2.714
California State Transportation Agency	2016	6.000
	2018	2.000
	2019	3.071
	2020	9.500
	2021	2.200
	2022	4.000
	2023	3.667
	2024	2.333
California State Water Resources Control Board	2020	27.222
	2021	13.062
	2022	5.625
	2023	7.500
	2024	9.000
California Strategic Growth Council	2016	4.000
	2017	5.200
	2018	5.400
	2019	8.667
	2020	18.143
	2021	4.500
	2022	3.000
California Wildlife Conservation Board	2022	19.000
	2023	6.500

1. **Evolution of Collaboration Models (2015-2024)**

- **Early Phase (2015-2017)**:
  - High DAC focus (75-100% benefit rate)
  - Smaller partnerships (2-4 counties)
  - Focus on equity over efficiency

- **Growth Phase (2018-2020)**:
  - Increased project size
  - Expanded partnerships
  - Attempt to balance equity and efficiency
- **Recent Phase (2021-2024)**:
  - Larger projects ($780K vs $462K pre-2020)
  - More partners (4.82 vs 3.73 pre-2020)
  - Focus shifted to efficiency over equity

2. **Regional Collaboration Success Models**

- **Northern California Success**:
  - Most projects (461)
  - Highest average funding ($805K)
  - Best GHG efficiency (0.006)
  - Strong DAC rate (10.8%)
- **Multi-Region Effectiveness**:
  - Larger partnerships (12.4 partners on average)
  - Higher DAC benefits (14.9%)
  - Good GHG efficiency (0.004)

3. **Agency Adaptation Patterns**

- **CalSTA's Evolution**:
  - Early: High DAC focus (100%)
  - Growth: Balanced approach (19.2% DAC, 0.017 GHG)
  - Recent: Efficiency focus (0.012 GHG)
- **CalFire's Transformation**:
  - Early: Strong equity (76.9% DAC)
  - Growth: Balanced (6.7% DAC, 0.007 GHG)
  - Recent: Efficiency focus (0.003 GHG)

4. **Policy Implications**

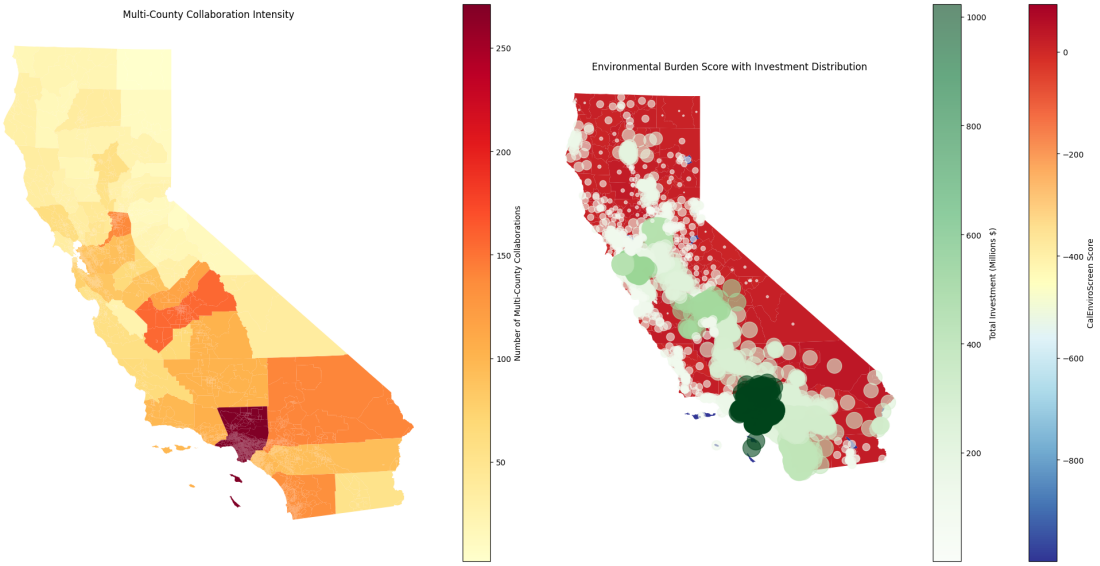
- Successful collaborations require:
  - Optimal partner size (2-8 counties)
  - Regional adjacency
  - Shared environmental challenges
  - Clear agency leadership

This analysis suggests that while collaboration capacity has grown, there's been a clear trade-off between equity and efficiency goals. The challenge for future policy is maintaining equity focus while scaling up projects.

[ ]:
```
import geopandas as gpd
import matplotlib.pyplot as plt

# First map: Collaboration Intensity
# Count how many times each county appears in collaborations
def count_county_collaborations(county_list):
    county_counts = {}
```

```
    for collab in multi_county['County']:
        counties = [c.strip() for c in str(collab).split(',')]
        for county in counties:
            county_counts[county] = county_counts.get(county, 0) + 1
    return county_counts

collab_counts = count_county_collaborations(gdf['County'].unique())

# Add collaboration counts to gdf
gdf['collab_count'] = gdf['County'].map(collab_counts).fillna(0)

# Create the first map
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(20, 10))

# Collaboration Intensity Map
gdf.plot(column='collab_count',
         ax=ax1,
         legend=True,
         legend_kwds={'label': 'Number of Multi-County Collaborations'},
         cmap='YlOrRd')
ax1.set_title('Multi-County Collaboration Intensity')
ax1.axis('off')

# Environmental Burden and Investment Map
# Add investment data to gdf
county_investments = data_filtered.groupby('County')['Total Program␣
 ↪GGRFFunding'].sum()/1e6  # Convert to millions
gdf['total_investment'] = gdf['County'].map(county_investments).fillna(0)

# Create second map
gdf.plot(column='CIscore',
         ax=ax2,
         legend=True,
         legend_kwds={'label': 'CalEnviroScreen Score'},
         cmap='RdYlBu_r')

# Add investment amount as points
scatter = ax2.scatter(gdf.geometry.centroid.x,
                      gdf.geometry.centroid.y,
                      c=gdf['total_investment'],
                      s=gdf['total_investment'].apply(lambda x: min(x*2, 500)), ␣
 ↪# Scale point size
                      cmap='Greens',
                      alpha=0.6)
ax2.set_title('Environmental Burden Score with Investment Distribution')
ax2.axis('off')
```

```
# Add colorbar for investment
plt.colorbar(scatter, ax=ax2, label='Total Investment (Millions $)')

plt.tight_layout()
plt.show()

# Print summary statistics
print("\nTop 10 Counties by Collaboration Count:")
print(pd.Series(collab_counts).sort_values(ascending=False).head(10))

print("\nCorrelation between CES Score and Investment:")
correlation = gdf['CIscore'].corr(gdf['total_investment'])
print(f"Correlation coefficient: {correlation:.3f}")
```




```
Top 10 Counties by Collaboration Count:
Los Angeles       271
Fresno            156
San Bernardino    140
Sacramento        138
San Diego         134
Madera            116
Ventura           104
Tulare            102
Kern              102
Orange             98
dtype: int64

Correlation between CES Score and Investment:
```

```
Correlation coefficient: 0.020
```