Matching Exercise

Nick Carroll & Yuanjing Zhu

In [ ]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import scipy.stats as stats
import altair as alt
%config InlineBackend.figure_format = 'retina'
import warnings
warnings.filterwarnings('ignore')
```

In [ ]:
```
cps = pd.read_stata(
    "https://github.com/nickeubank/MIDS_Data/blob/master"
    "/Current_Population_Survey/cps_for_matching.dta?raw=true"
)
```

In [ ]:
```
cps.head()
```

Out[ ]:

	index	annual_earnings	female	simplified_race	has_college	age	county	class94
0	151404	NaN	1	3.0	1	30	0-WV	Private, For Profit
1	123453	NaN	0	0.0	0	21	251-TX	Private, For Profit
2	187982	NaN	0	0.0	0	40	5-MA	Self-Employed, Unincorporated
3	122356	NaN	1	0.0	1	27	0-TN	Private, Nonprofit
4	210750	42900.0	1	0.0	0	52	0-IA	Private, For Profit

In [ ]:
```
cps.loc[:, 'county']
```

Out[ ]:
```
0          0-WV
1        251-TX
2          5-MA
3          0-TN
4          0-IA
          ...
11145     13-AZ
11146     35-NJ
11147      0-MS
11148      1-DC
11149     15-NH
Name: county, Length: 11150, dtype: object
```

Exercise 1: raw difference of annual_earnings between those with and without a college degree

In [ ]:
```
annual_earnings_with_college = cps.loc[cps['has_college']==1, 'annual_earnings']
annual_earnings_without_college = cps.loc[cps['has_college']==0, 'annual_earnings']

raw_diff = annual_earnings_with_college.mean() - annual_earnings_without_college.mean()
print(f'The raw difference of annual_earnings between those with and without a college degree is ${raw_diff:.2f}')
```

```
The raw difference of annual_earnings between those with and without a college degree is $14158.50
```

In [ ]:
```
_, pvalue = stats.ttest_ind(annual_earnings_with_college, annual_earnings_without_college, nan_policy='omit')
print(f'The p-value for the t-test is {pvalue:.2f}')
```

```
The p-value for the t-test is 0.00
```

The difference is statistically significant

Exercise 2: check for balance

- **Race**

In [ ]:
```
cps['simplified_race'].value_counts()
```

Out[ ]:
```
0.0    7622
2.0    1512
1.0    1020
3.0     996
Name: simplified_race, dtype: int64
```

Race is coded as White Non-Hispanic (0), Black Non-Hispanic (1), Hispanic (2), Other (3).

In [ ]:
```
college_by_race = cps.groupby("simplified_race")["has_college"].mean()

print("Share of people in different racial groups who have college degrees:")
print(college_by_race)
```

```
Share of people in different racial groups who have college degrees:
simplified_race
0.0    0.438205
1.0    0.317647
2.0    0.198413
3.0    0.474900
Name: has_college, dtype: float64
```

In [ ]:
```python
cross_tab = pd.crosstab(cps['simplified_race'], cps['has_college'])
_, p, _, _ = stats.chi2_contingency(cross_tab)
print(f"P-value for chi-square test: {p:.2f}")
```

```
P-value for chi-square test: 0.00
```

Those difference in the share of people in different racial groups who have college degrees is statistically significant

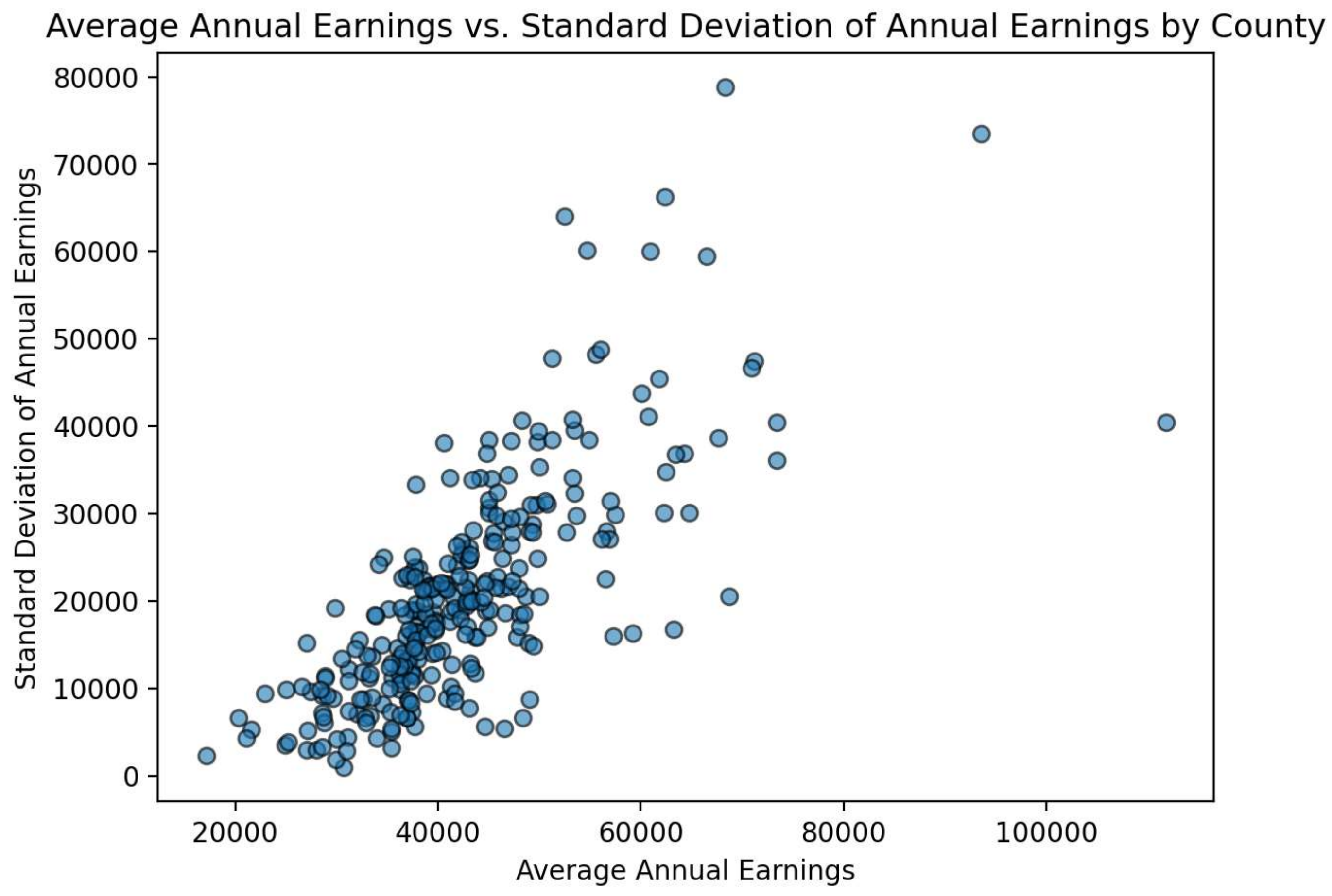
- **County**

In [ ]:
```python
cps.groupby('county')['has_college'].mean()
```

Out[ ]:
```
county
0-AK     0.323308
0-AL     0.337143
0-AR     0.298246
0-AZ     0.714286
0-CA     0.395455
           ...
97-IL    0.520000
99-CA    0.133333
99-FL    0.583333
99-MI    0.476190
99-MO    0.333333
Name: has_college, Length: 326, dtype: float64
```

In [ ]:
```python
# plot the difference
l_avg = []
l_std = []
for i in cps['county'].unique():
    l_avg.append(cps.loc[cps['county']==i, 'annual_earnings'].mean())
    l_std.append(cps.loc[cps['county']==i, 'annual_earnings'].std())

plt.figure(figsize=(7,5), dpi=100)
plt.scatter(l_avg, l_std, edgecolors='black', alpha=0.6)
plt.xlabel('Average Annual Earnings')
plt.ylabel('Standard Deviation of Annual Earnings')
plt.title('Average Annual Earnings vs. Standard Deviation of Annual Earnings by County')
plt.show()
```



The distribution the share of people who have college degrees looks different across counties

In terms of race and county, the data seems to be imbalanced.

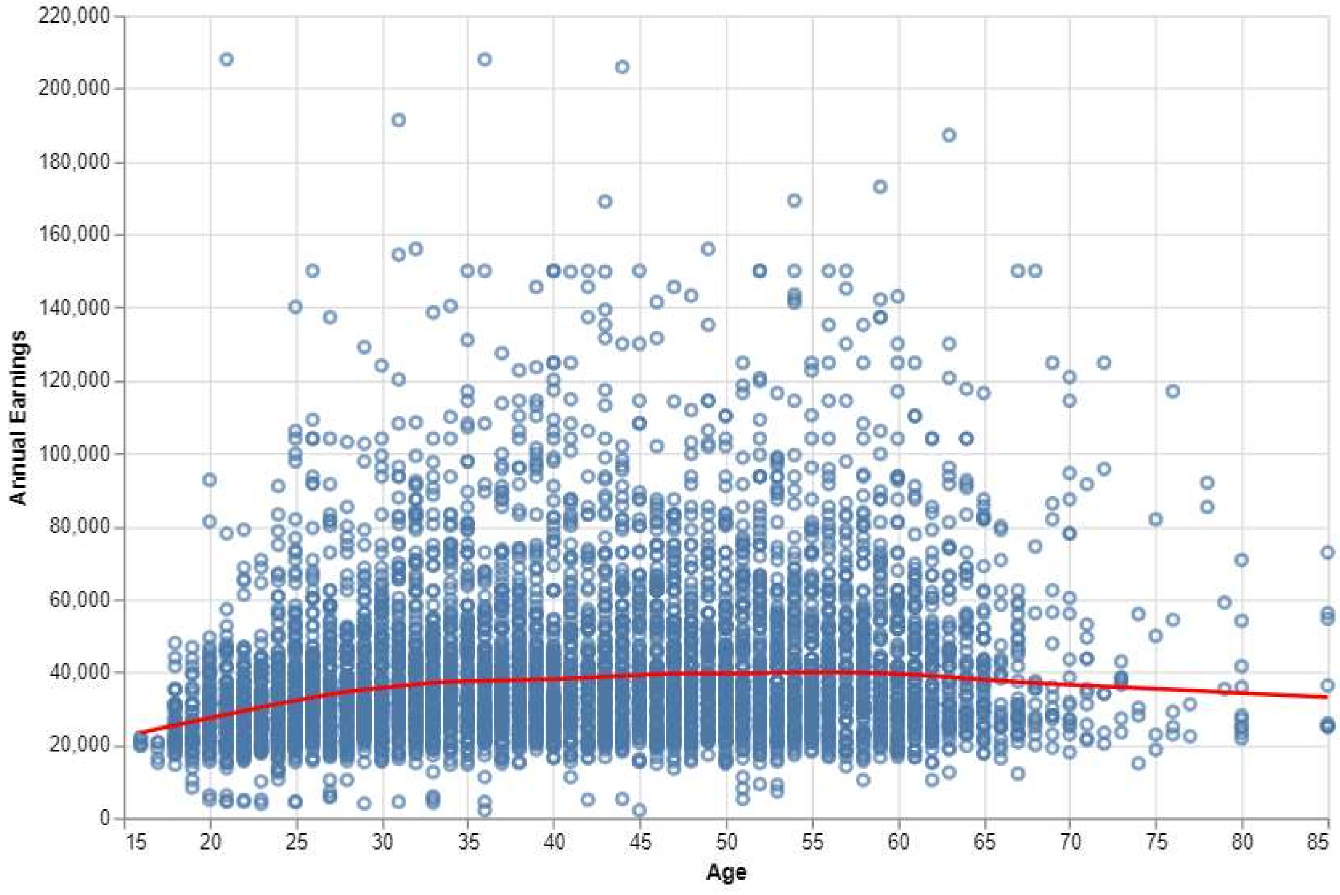
Exercise 3: Plot a non-linear regression of annual_earnings on age

In [ ]:
```python
alt.data_transformers.disable_max_rows()
alt.renderers.set_embed_options(width=600, height=400)
base = (
    alt.Chart(cps).mark_point().encode(
        x = alt.X('age', title='Age', scale=alt.Scale(zero=False)),
        y = alt.Y('annual_earnings',title='Annual Earnings', scale=alt.Scale(zero=False)),
    )
)
# change font size
base.configure_title(fontSize=20)
```

```
# fit a regression line
loess = base.transform_loess(
    'age', 'annual_earnings').mark_line(color='red')

base + loess
```

Out[ ]:



- The relationship between age and annual earnings looks non-linear.
- Since there is no clear relationship between age and earnings, using matching allows us to compare people who are the same age (or a representative match) so that the actual relationship between age and earnings doesn't matter.

Exercise 4: Create a new variable that discretizes age into a single value for each decade of age.

In [ ]:
```
print(f"Age range: {cps['age'].min()} to {cps['age'].max()}")
```

```
Age range: 16 to 85
```

In [ ]:
```
bins = [0, 20, 30, 40, 50, 60, 70, 80, 90]
pd.cut(cps['age'], bins=bins).value_counts().sort_index()
```

Out[ ]:
```
(0, 20]      214
(20, 30]    2047
(30, 40]    2737
(40, 50]    2582
(50, 60]    2315
(60, 70]    1082
(70, 80]     156
(80, 90]      17
Name: age, dtype: int64
```

Create a new variable that discretizes age:

- 0-30: 0
- 31-40: 1
- 41-50: 2
- 51-60: 3
- 61-70: 4
- 71-90: 5

In [ ]:
```
# create a new varibalbe that discretizes age into 7 bins
cps.loc[cps['age'] < 30, 'age_group'] = 0
cps.loc[(cps['age'] >= 30) & (cps['age'] < 40), 'age_group'] = 1
cps.loc[(cps['age'] >= 40) & (cps['age'] < 50), 'age_group'] = 2
cps.loc[(cps['age'] >= 50) & (cps['age'] < 60), 'age_group'] = 3
cps.loc[(cps['age'] >= 60) & (cps['age'] < 70), 'age_group'] = 4
cps.loc[cps['age'] >= 70, 'age_group'] = 5
cps['age_group'] = cps['age_group'].astype("category")
```

Exercise 5: covert our string variables into numeric variables for DAME

In [ ]:
```
# convert county and class94 to a numeric vector of intergers using pd.Categorical
cps['county'] = pd.Categorical(cps['county']).codes
cps['class94'] = pd.Categorical(cps['class94']).codes
```

Exercise 6: drop all the variables we don’t need

In [ ]:
```
# drop the original variables that we dont need
cps_2 = cps.copy()    # create a copy of the dataframe
cps_2.drop(columns=['age', 'index'], inplace=True)
cps_2.dropna(inplace=True)
cps_2.reset_index(inplace=True, drop=True)
cps_2.head()
```

Out[ ]:

	annual_earnings	female	simplified_race	has_college	county	class94	age_group
0	42900.0	1	0.0	0	10	3	3.0
1	31200.0	0	2.0	0	31	3	1.0
2	20020.0	0	0.0	1	8	3	4.0
3	22859.2	0	0.0	0	44	1	2.0
4	73860.8	0	0.0	1	24	3	1.0

Exercise 7: Start matching

In [ ]:

```
import dame_flame
model = dame_flame.matching.DAME(repeats=False, verbose=3, want_pe=True)
model.fit(
    cps_2,
    treatment_column_name="has_college",
    outcome_column_name="annual_earnings",
)
result = model.predict(cps_2)
```

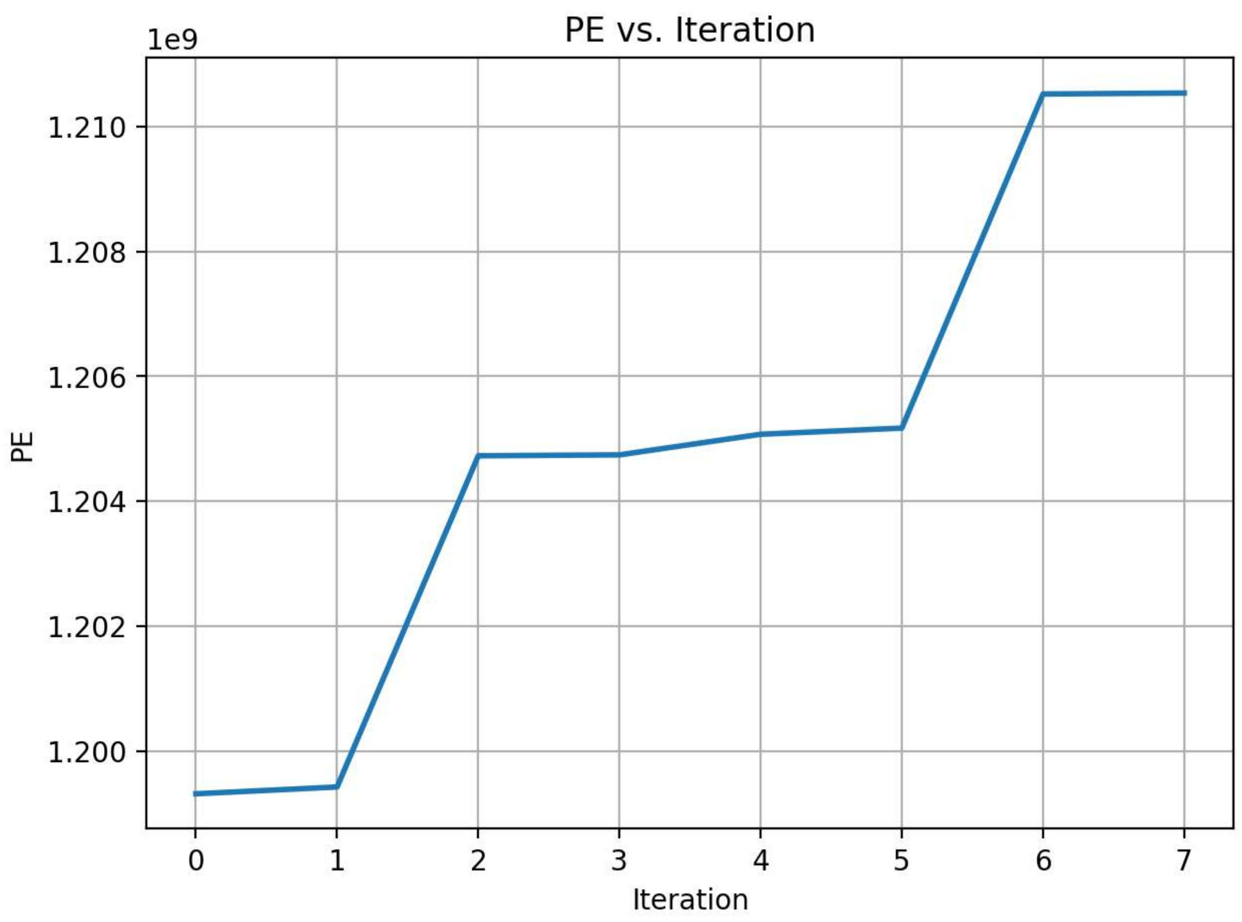
```
Completed iteration 0 of matching
	Number of matched groups formed in total:  370
	Unmatched treated units:  644 out of a total of  1150 treated units
	Unmatched control units:  3187 out of a total of  4365 control units
	Number of matches made this iteration:  1684
	Number of matches made so far:  1684
	Covariates dropped so far:  set()
	Predictive error of covariate set used to match:  1199312680.0957854
Completed iteration 1 of matching
	Number of matched groups formed in total:  494
	Unmatched treated units:  25 out of a total of  1150 treated units
	Unmatched control units:  180 out of a total of  4365 control units
	Number of matches made this iteration:  3626
	Number of matches made so far:  5310
	Covariates dropped so far:  frozenset({'county'})
	Predictive error of covariate set used to match:  1199421883.1095908
Completed iteration 2 of matching
	Number of matched groups formed in total:  494
	Unmatched treated units:  25 out of a total of  1150 treated units
	Unmatched control units:  180 out of a total of  4365 control units
	Number of matches made this iteration:  0
	Number of matches made so far:  5310
	Covariates dropped so far:  frozenset({'simplified_race'})
	Predictive error of covariate set used to match:  1204727749.8949614
Completed iteration 3 of matching
	Number of matched groups formed in total:  505
	Unmatched treated units:  8 out of a total of  1150 treated units
	Unmatched control units:  129 out of a total of  4365 control units
	Number of matches made this iteration:  68
	Number of matches made so far:  5378
	Covariates dropped so far:  frozenset({'simplified_race', 'county'})
	Predictive error of covariate set used to match:  1204742613.4791539
Completed iteration 4 of matching
	Number of matched groups formed in total:  505
	Unmatched treated units:  8 out of a total of  1150 treated units
	Unmatched control units:  129 out of a total of  4365 control units
	Number of matches made this iteration:  0
	Number of matches made so far:  5378
	Covariates dropped so far:  frozenset({'class94'})
	Predictive error of covariate set used to match:  1205072671.32629
Completed iteration 5 of matching
	Number of matched groups formed in total:  508
	Unmatched treated units:  5 out of a total of  1150 treated units
	Unmatched control units:  120 out of a total of  4365 control units
	Number of matches made this iteration:  12
	Number of matches made so far:  5390
	Covariates dropped so far:  frozenset({'class94', 'county'})
	Predictive error of covariate set used to match:  1205171280.4727237
Completed iteration 6 of matching
	Number of matched groups formed in total:  509
	Unmatched treated units:  4 out of a total of  1150 treated units
	Unmatched control units:  119 out of a total of  4365 control units
	Number of matches made this iteration:  2
	Number of matches made so far:  5392
	Covariates dropped so far:  frozenset({'class94', 'simplified_race'})
	Predictive error of covariate set used to match:  1210524158.7436352
Completed iteration 7 of matching
	Number of matched groups formed in total:  511
	Unmatched treated units:  0 out of a total of  1150 treated units
	Unmatched control units:  110 out of a total of  4365 control units
	Number of matches made this iteration:  13
	Number of matches made so far:  5405
	Covariates dropped so far:  frozenset({'class94', 'simplified_race', 'county'})
	Predictive error of covariate set used to match:  1210539313.933855
5405 units matched. We finished with no more treated units to match
```

Exercise 8: plot predicted errors in each iteration

In [ ]:

```
# plot pe vs. iteration
plt.figure(figsize=(7,5), dpi=100)
plt.plot(model.pe_each_iter, linewidth=2)
```

```
plt.grid('on')
plt.xlabel('Iteration')
plt.ylabel('PE')
plt.title('PE vs. Iteration')
plt.show()
```



From iteration 1 to iteration 2, and iteration 5 to iteration6, the matching quality seems to fall off dramatically.

Exercise 9: where to cut off the data?

In completed iteration 1 of matching, number of matches made so far is 5310, which is larger than 5000. In addtion, the match quality drops dramatically from iteration 1 to iteration 2. Therefore, we might **stop at iteration 1**.

Exercise 10: rerun matching using early_stop_iterations

In [ ]:
```
model_2 = dame_flame.matching.DAME(repeats=False, verbose=3, want_pe=True, early_stop_iterations=1)
model_2.fit(
    cps_2,
    treatment_column_name="has_college",
    outcome_column_name="annual_earnings",
)
result_2 = model_2.predict(cps_2)
```

```
Completed iteration 0 of matching
        Number of matched groups formed in total:  370
        Unmatched treated units:  644 out of a total of  1150 treated units
        Unmatched control units:  3187 out of a total of  4365 control units
        Number of matches made this iteration:  1684
        Number of matches made so far:  1684
        Covariates dropped so far:  set()
        Predictive error of covariate set used to match:  1199312680.0957854
Completed iteration 1 of matching
        Number of matched groups formed in total:  494
        Unmatched treated units:  25 out of a total of  1150 treated units
        Unmatched control units:  180 out of a total of  4365 control units
        Number of matches made this iteration:  3626
        Number of matches made so far:  5310
        Covariates dropped so far:  frozenset({'county'})
        Predictive error of covariate set used to match:  1199421883.1095908
5310 units matched. We stopped after iteration 1
```

Exercise 11: get back dataset

In [ ]:
```
def get_dataframe(model, result_of_fit):

    # Get original data
    better = model.input_data.loc[result_of_fit.index]
    if not better.index.is_unique:
        raise ValueError("Need index values in input data to be unique")

    # Get match groups for clustering
    better["match_group"] = np.nan
    better["match_group_size"] = np.nan
    for idx, group in enumerate(model.units_per_group):
        better.loc[group, "match_group"] = idx
        better.loc[group, "match_group_size"] = len(group)

    # Get weights. I THINK this is right?! At least for with repeat=False?
    t = model.treatment_column_name
    better["t_in_group"] = better.groupby("match_group")[t].transform(np.sum)

    # Make weights
    better["weights"] = np.nan
    better.loc[better[t] == 1, "weights"] = 1  # treaments are 1
```

```python
    # Controls start as proportional to num of treatments
    # each observation is matched to.
    better.loc[better[t] == 0, "weights"] = better["t_in_group"] / (
        better["match_group_size"] - better["t_in_group"]
    )

    # Then re-normalize for num unique control observations.
    control_weights = better[better[t] == 0]["weights"].sum()

    num_control_obs = len(better[better[t] == 0].index.drop_duplicates())
    renormalization = num_control_obs / control_weights
    better.loc[better[t] == 0, "weights"] = (
        better.loc[better[t] == 0, "weights"] * renormalization
    )
    assert better.weights.notnull().all()

    better = better.drop(["t_in_group"], axis="columns")

    # Make sure right length and values!
    assert len(result_of_fit) == len(better)
    assert better.loc[better[t] == 0, "weights"].sum() == num_control_obs

    return better
```

In [ ]:
```python
matched_data = get_dataframe(model_2, result_2)
matched_data.head()
```

Out[ ]:

	annual_earnings	female	simplified_race	has_college	county	class94	age_group	match_group	match_group_size	weights
0	42900.0	1	0.0	0	10	3	3.0	59.0	5.0	0.930000
1	31200.0	0	2.0	0	31	3	1.0	411.0	108.0	0.070189
2	20020.0	0	0.0	1	8	3	4.0	52.0	3.0	1.000000
3	22859.2	0	0.0	0	44	1	2.0	424.0	28.0	1.240000
4	73860.8	0	0.0	1	24	3	1.0	106.0	7.0	1.000000

Exercise 12: check balance again

In [ ]:
```python
import statsmodels.formula.api as smf
```

In [ ]:
```python
smf.wls(
    "has_college ~ C(simplified_race)", matched_data, weights=matched_data["weights"]
).fit().summary()
```

Out[ ]:

WLS Regression Results

Dep. Variable:	has_college	**R-squared:**	0.000
Model:	WLS	**Adj. R-squared:**	-0.001
Method:	Least Squares	**F-statistic:**	2.268e-13
Date:	Wed, 01 Mar 2023	**Prob (F-statistic):**	1.00
Time:	22:48:44	**Log-Likelihood:**	-3736.0
No. Observations:	5310	**AIC:**	7480.
Df Residuals:	5306	**BIC:**	7506.
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	0.2119	0.007	31.608	0.000	0.199	0.225
C(simplified_race)[T.1.0]	-3.209e-17	0.018	-1.77e-15	1.000	-0.036	0.036
C(simplified_race)[T.2.0]	-1.431e-16	0.019	-7.62e-15	1.000	-0.037	0.037
C(simplified_race)[T.3.0]	-5.031e-17	0.020	-2.48e-15	1.000	-0.040	0.040

Omnibus:	860.389	**Durbin-Watson:**	2.000
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	1353.227
Skew:	1.234	**Prob(JB):**	1.41e-294
Kurtosis:	2.851	**Cond. No.**	3.95

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [ ]:
```python
smf.wls(
    "annual_earnings ~ C(county)", matched_data, weights=matched_data["weights"]
).fit().summary()
```

Out[ ]:

WLS Regression Results

Dep. Variable:	annual_earnings	**R-squared:**	0.097
Model:	WLS	**Adj. R-squared:**	0.040
Method:	Least Squares	**F-statistic:**	1.704
Date:	Wed, 01 Mar 2023	**Prob (F-statistic):**	8.82e-13
Time:	22:48:44	**Log-Likelihood:**	-61639.
No. Observations:	5310	**AIC:**	1.239e+05
Df Residuals:	4994	**BIC:**	1.260e+05
Df Model:	315		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	5.244e+04	2507.026	20.916	0.000	4.75e+04	5.74e+04
C(county)[T.1]	-1.189e+04	3488.641	-3.408	0.001	-1.87e+04	-5049.610
C(county)[T.2]	-1.469e+04	3503.012	-4.195	0.000	-2.16e+04	-7826.209
C(county)[T.3]	1304.5895	1.28e+04	0.102	0.919	-2.39e+04	2.65e+04
C(county)[T.4]	-8760.4917	3219.364	-2.721	0.007	-1.51e+04	-2449.124
C(county)[T.5]	-6906.6690	4527.680	-1.525	0.127	-1.58e+04	1969.572
C(county)[T.6]	-1.643e+04	1.13e+04	-1.455	0.146	-3.86e+04	5705.827
C(county)[T.7]	-9093.7230	4147.822	-2.192	0.028	-1.72e+04	-962.171
C(county)[T.8]	-1.389e+04	3416.382	-4.066	0.000	-2.06e+04	-7192.779
C(county)[T.9]	-1.361e+04	5569.268	-2.444	0.015	-2.45e+04	-2691.879
C(county)[T.10]	-1.579e+04	3708.085	-4.258	0.000	-2.31e+04	-8520.931
C(county)[T.11]	-1.251e+04	3283.832	-3.808	0.000	-1.89e+04	-6068.288
C(county)[T.12]	-6481.0973	3380.473	-1.917	0.055	-1.31e+04	146.113
C(county)[T.13]	-1.372e+04	3642.260	-3.766	0.000	-2.09e+04	-6576.261
C(county)[T.14]	-1.542e+04	4244.139	-3.632	0.000	-2.37e+04	-7095.967
C(county)[T.15]	-9856.6109	5203.410	-1.894	0.058	-2.01e+04	344.359
C(county)[T.16]	-1.122e+04	3907.170	-2.870	0.004	-1.89e+04	-3555.611
C(county)[T.17]	-1.357e+04	8393.133	-1.617	0.106	-3e+04	2881.894
C(county)[T.18]	-1503.2836	5153.599	-0.292	0.771	-1.16e+04	8600.034
C(county)[T.19]	-1.474e+04	7114.322	-2.072	0.038	-2.87e+04	-793.047
C(county)[T.20]	-1.172e+04	4693.457	-2.497	0.013	-2.09e+04	-2517.918
C(county)[T.21]	-3219.1965	3877.316	-0.830	0.406	-1.08e+04	4382.045
C(county)[T.22]	-1.743e+04	3993.574	-4.364	0.000	-2.53e+04	-9597.991
C(county)[T.23]	-1.502e+04	3483.642	-4.312	0.000	-2.19e+04	-8192.516
C(county)[T.24]	-9497.6392	3427.725	-2.771	0.006	-1.62e+04	-2777.792
C(county)[T.25]	-1.939e+04	3769.434	-5.143	0.000	-2.68e+04	-1.2e+04
C(county)[T.26]	-1.451e+04	3395.786	-4.273	0.000	-2.12e+04	-7851.875
C(county)[T.27]	-1.306e+04	3618.312	-3.609	0.000	-2.02e+04	-5963.829
C(county)[T.28]	-1.336e+04	6158.954	-2.169	0.030	-2.54e+04	-1282.285
C(county)[T.29]	-6847.5608	5646.561	-1.213	0.225	-1.79e+04	4222.179
C(county)[T.30]	-1078.0211	5467.386	-0.197	0.844	-1.18e+04	9640.457
C(county)[T.31]	-1.271e+04	5112.457	-2.486	0.013	-2.27e+04	-2684.603
C(county)[T.32]	-5644.1071	3936.402	-1.434	0.152	-1.34e+04	2072.969
C(county)[T.33]	-1.166e+04	3313.565	-3.518	0.000	-1.82e+04	-5162.139
C(county)[T.34]	-1.207e+04	3739.178	-3.229	0.001	-1.94e+04	-4744.562
C(county)[T.35]	-1.228e+04	3664.741	-3.350	0.001	-1.95e+04	-5092.899
C(county)[T.36]	-1.016e+04	3950.027	-2.572	0.010	-1.79e+04	-2414.984
C(county)[T.37]	-1.247e+04	4078.440	-3.057	0.002	-2.05e+04	-4471.027
C(county)[T.38]	-1.496e+04	3565.662	-4.194	0.000	-2.19e+04	-7964.861
C(county)[T.39]	-1.121e+04	3525.817	-3.178	0.001	-1.81e+04	-4294.476
C(county)[T.40]	-1.44e+04	3983.539	-3.616	0.000	-2.22e+04	-6594.207
C(county)[T.41]	-1.098e+04	2892.228	-3.798	0.000	-1.67e+04	-5314.525
C(county)[T.42]	-1.284e+04	3739.701	-3.434	0.001	-2.02e+04	-5511.965
C(county)[T.43]	-1.021e+04	4001.653	-2.552	0.011	-1.81e+04	-2368.546
C(county)[T.44]	-8071.0611	3198.617	-2.523	0.012	-1.43e+04	-1800.367
C(county)[T.45]	-5966.4238	3385.006	-1.763	0.078	-1.26e+04	669.674
C(county)[T.46]	-8582.8658	3659.596	-2.345	0.019	-1.58e+04	-1408.450

C(county)[T.47]	-1.5e+04	3257.097	-4.605	0.000	-2.14e+04	-8614.014
C(county)[T.48]	-8609.2032	3567.059	-2.414	0.016	-1.56e+04	-1616.201
C(county)[T.49]	2.323e+04	8667.367	2.681	0.007	6242.144	4.02e+04
C(county)[T.50]	-2.085e+04	1.22e+04	-1.703	0.089	-4.49e+04	3153.374
C(county)[T.51]	-9576.4201	3663.554	-2.614	0.009	-1.68e+04	-2394.246
C(county)[T.52]	-1.782e+04	8999.837	-1.981	0.048	-3.55e+04	-180.975
C(county)[T.53]	-2345.4585	2.14e+04	-0.110	0.913	-4.42e+04	3.96e+04
C(county)[T.54]	-1.607e+04	9617.853	-1.670	0.095	-3.49e+04	2789.359
C(county)[T.55]	8521.2597	1.05e+04	0.810	0.418	-1.21e+04	2.92e+04
C(county)[T.56]	-9783.1290	7887.849	-1.240	0.215	-2.52e+04	5680.519
C(county)[T.57]	1.375e+04	1.52e+04	0.906	0.365	-1.6e+04	4.35e+04
C(county)[T.58]	-5989.5416	5705.690	-1.050	0.294	-1.72e+04	5196.116
C(county)[T.59]	-1.846e+04	1.07e+04	-1.720	0.086	-3.95e+04	2581.376
C(county)[T.60]	-1.731e+04	7282.390	-2.376	0.018	-3.16e+04	-3029.557
C(county)[T.61]	-5882.5789	1.19e+04	-0.495	0.621	-2.92e+04	1.74e+04
C(county)[T.62]	1.538e+04	1.44e+04	1.069	0.285	-1.28e+04	4.36e+04
C(county)[T.63]	-1.283e+04	9362.422	-1.370	0.171	-3.12e+04	5525.240
C(county)[T.64]	1.576e+04	9326.736	1.690	0.091	-2521.646	3.4e+04
C(county)[T.65]	-1.988e+04	9966.396	-1.994	0.046	-3.94e+04	-338.467
C(county)[T.66]	-2.177e+04	1.6e+04	-1.358	0.174	-5.32e+04	9648.277
C(county)[T.67]	-1.277e+04	3.22e+04	-0.396	0.692	-7.6e+04	5.04e+04
C(county)[T.68]	3826.6775	1.04e+04	0.368	0.713	-1.66e+04	2.42e+04
C(county)[T.69]	-1.553e+04	9824.699	-1.581	0.114	-3.48e+04	3731.454
C(county)[T.70]	-1.795e+04	1.45e+04	-1.242	0.214	-4.63e+04	1.04e+04
C(county)[T.71]	-1.76e+04	9777.817	-1.800	0.072	-3.68e+04	1563.926
C(county)[T.72]	1.972e+04	2.07e+04	0.950	0.342	-2.1e+04	6.04e+04
C(county)[T.73]	-1877.1650	6113.906	-0.307	0.759	-1.39e+04	1.01e+04
C(county)[T.74]	-1.674e+04	8176.546	-2.047	0.041	-3.28e+04	-710.380
C(county)[T.75]	-1.65e+04	4245.321	-3.887	0.000	-2.48e+04	-8177.711
C(county)[T.76]	-3.385e+04	2.85e+04	-1.189	0.235	-8.97e+04	2.2e+04
C(county)[T.77]	-6909.4702	7439.851	-0.929	0.353	-2.15e+04	7675.905
C(county)[T.78]	-957.2981	2.43e+04	-0.039	0.969	-4.87e+04	4.67e+04
C(county)[T.79]	3.436e+04	9182.599	3.742	0.000	1.64e+04	5.24e+04
C(county)[T.80]	-1.931e+04	6059.695	-3.187	0.001	-3.12e+04	-7429.590
C(county)[T.81]	-1.095e+04	6198.630	-1.767	0.077	-2.31e+04	1198.645
C(county)[T.82]	-1.662e+04	1.75e+04	-0.951	0.342	-5.09e+04	1.76e+04
C(county)[T.83]	-2.405e+04	1.36e+04	-1.766	0.077	-5.07e+04	2645.345
C(county)[T.84]	-1.284e+04	1.32e+04	-0.971	0.332	-3.88e+04	1.31e+04
C(county)[T.85]	8310.4618	9819.165	0.846	0.397	-1.09e+04	2.76e+04
C(county)[T.86]	3.117e+04	1.55e+04	2.012	0.044	795.473	6.15e+04
C(county)[T.87]	-2.31e+04	1.18e+04	-1.952	0.051	-4.63e+04	101.213
C(county)[T.88]	-9128.9205	1.89e+04	-0.484	0.629	-4.61e+04	2.79e+04
C(county)[T.89]	-1.861e+04	9258.807	-2.009	0.045	-3.68e+04	-454.123
C(county)[T.90]	-9054.6425	5382.947	-1.682	0.093	-1.96e+04	1498.298
C(county)[T.91]	-7944.4596	1.06e+04	-0.748	0.454	-2.88e+04	1.29e+04
C(county)[T.92]	-1.388e+04	1.27e+04	-1.092	0.275	-3.88e+04	1.1e+04
C(county)[T.93]	-6270.6798	9821.740	-0.638	0.523	-2.55e+04	1.3e+04
C(county)[T.94]	2869.9066	1.06e+04	0.271	0.786	-1.79e+04	2.36e+04
C(county)[T.95]	-1.116e+04	5469.587	-2.041	0.041	-2.19e+04	-439.635
C(county)[T.96]	-9626.7558	1.12e+04	-0.858	0.391	-3.16e+04	1.24e+04
C(county)[T.97]	-2.544e+04	1.02e+04	-2.487	0.013	-4.55e+04	-5383.249
C(county)[T.98]	3868.2644	1.66e+04	0.232	0.816	-2.87e+04	3.65e+04
C(county)[T.99]	-9579.5366	3351.117	-2.859	0.004	-1.61e+04	-3009.875
C(county)[T.100]	403.7017	8517.201	0.047	0.962	-1.63e+04	1.71e+04
C(county)[T.101]	-1.294e+04	1.04e+04	-1.241	0.215	-3.34e+04	7505.528
C(county)[T.102]	-9343.0656	1.47e+04	-0.637	0.524	-3.81e+04	1.94e+04
C(county)[T.103]	-8667.4236	9845.095	-0.880	0.379	-2.8e+04	1.06e+04
C(county)[T.104]	-669.4598	1.26e+04	-0.053	0.958	-2.54e+04	2.41e+04

C(county)[T.105]	-8942.8456	1.41e+04	-0.633	0.527	-3.67e+04	1.88e+04
C(county)[T.106]	-3.164e+04	3.64e+04	-0.868	0.385	-1.03e+05	3.98e+04
C(county)[T.107]	-2.086e+04	1.86e+04	-1.120	0.263	-5.74e+04	1.56e+04
C(county)[T.109]	-1036.6820	8707.780	-0.119	0.905	-1.81e+04	1.6e+04
C(county)[T.110]	-1.062e+04	1.15e+04	-0.923	0.356	-3.32e+04	1.19e+04
C(county)[T.111]	-2.035e+04	2.38e+04	-0.855	0.392	-6.7e+04	2.63e+04
C(county)[T.112]	-6504.1746	1.55e+04	-0.420	0.674	-3.68e+04	2.38e+04
C(county)[T.113]	-1.884e+04	2.83e+04	-0.666	0.505	-7.42e+04	3.66e+04
C(county)[T.114]	-2.475e+04	1.75e+04	-1.414	0.157	-5.91e+04	9560.670
C(county)[T.115]	7838.4225	1.08e+04	0.722	0.470	-1.34e+04	2.91e+04
C(county)[T.116]	-1.6e+04	1.47e+04	-1.090	0.276	-4.48e+04	1.28e+04
C(county)[T.117]	-2.956e+04	4.66e+04	-0.635	0.526	-1.21e+05	6.17e+04
C(county)[T.118]	-1.782e+04	1.2e+04	-1.485	0.138	-4.13e+04	5703.042
C(county)[T.120]	-1.664e+04	1.44e+04	-1.157	0.247	-4.48e+04	1.16e+04
C(county)[T.121]	-1.462e+04	1.44e+04	-1.018	0.309	-4.28e+04	1.35e+04
C(county)[T.123]	-7023.6920	5512.814	-1.274	0.203	-1.78e+04	3783.845
C(county)[T.124]	-2.729e+04	1.09e+04	-2.506	0.012	-4.86e+04	-5942.911
C(county)[T.125]	-5151.9648	1.35e+04	-0.382	0.702	-3.16e+04	2.13e+04
C(county)[T.126]	-2.115e+04	1.06e+04	-1.993	0.046	-4.2e+04	-350.362
C(county)[T.127]	-7747.4395	1.04e+04	-0.746	0.456	-2.81e+04	1.26e+04
C(county)[T.128]	-1.032e+04	3.64e+04	-0.283	0.777	-8.18e+04	6.11e+04
C(county)[T.129]	-1.212e+04	1.33e+04	-0.914	0.361	-3.81e+04	1.39e+04
C(county)[T.130]	-2.887e+04	1.35e+04	-2.137	0.033	-5.54e+04	-2389.667
C(county)[T.131]	-4337.2981	2.31e+04	-0.188	0.851	-4.96e+04	4.09e+04
C(county)[T.132]	-4744.8556	6111.516	-0.776	0.438	-1.67e+04	7236.399
C(county)[T.133]	1.1e+04	1.64e+04	0.671	0.503	-2.12e+04	4.32e+04
C(county)[T.134]	6.304e+04	1.95e+04	3.235	0.001	2.48e+04	1.01e+05
C(county)[T.135]	-5678.4076	1.5e+04	-0.379	0.705	-3.51e+04	2.37e+04
C(county)[T.136]	-8670.3207	6499.263	-1.334	0.182	-2.14e+04	4071.089
C(county)[T.137]	-1.549e+04	8773.359	-1.766	0.078	-3.27e+04	1708.337
C(county)[T.138]	-6706.0083	8720.461	-0.769	0.442	-2.38e+04	1.04e+04
C(county)[T.139]	-2241.9792	1.11e+04	-0.201	0.840	-2.41e+04	1.96e+04
C(county)[T.140]	-2.118e+04	1.49e+04	-1.419	0.156	-5.04e+04	8086.948
C(county)[T.141]	-7657.1257	1.33e+04	-0.577	0.564	-3.37e+04	1.84e+04
C(county)[T.142]	7622.7019	2.31e+04	0.330	0.741	-3.76e+04	5.28e+04
C(county)[T.143]	-7150.6250	6364.153	-1.124	0.261	-1.96e+04	5325.910
C(county)[T.144]	-2.633e+04	2.31e+04	-1.141	0.254	-7.16e+04	1.89e+04
C(county)[T.145]	-1726.8981	3.77e+04	-0.046	0.963	-7.56e+04	7.21e+04
C(county)[T.146]	-1.511e+04	1.73e+04	-0.873	0.382	-4.9e+04	1.88e+04
C(county)[T.147]	-2.363e+04	1.64e+04	-1.440	0.150	-5.58e+04	8538.604
C(county)[T.148]	-3.01e+04	1.9e+04	-1.580	0.114	-6.74e+04	7240.768
C(county)[T.149]	-1.533e+04	5192.577	-2.953	0.003	-2.55e+04	-5153.506
C(county)[T.150]	-1.677e+04	3.64e+04	-0.460	0.646	-8.82e+04	5.47e+04
C(county)[T.151]	-2.271e+04	1.23e+04	-1.846	0.065	-4.68e+04	1406.433
C(county)[T.152]	-2.095e+04	7928.367	-2.642	0.008	-3.65e+04	-5407.282
C(county)[T.153]	-1.5e+04	3.24e+04	-0.462	0.644	-7.86e+04	4.86e+04
C(county)[T.155]	-1.092e+04	1.43e+04	-0.763	0.445	-3.9e+04	1.71e+04
C(county)[T.157]	-1.954e+04	1.45e+04	-1.344	0.179	-4.8e+04	8953.121
C(county)[T.158]	-3.03e+04	1.34e+04	-2.257	0.024	-5.66e+04	-3978.775
C(county)[T.159]	-2436.9330	2.9e+04	-0.084	0.933	-5.93e+04	5.44e+04
C(county)[T.160]	-1.838e+04	9057.956	-2.030	0.042	-3.61e+04	-626.752
C(county)[T.161]	9252.1731	9290.483	0.996	0.319	-8961.253	2.75e+04
C(county)[T.162]	9286.6263	1.68e+04	0.552	0.581	-2.37e+04	4.23e+04
C(county)[T.163]	-2.966e+04	2.2e+04	-1.349	0.177	-7.28e+04	1.35e+04
C(county)[T.164]	-9797.2981	1.64e+04	-0.597	0.550	-4.2e+04	2.24e+04
C(county)[T.165]	-1.635e+04	2.32e+04	-0.706	0.481	-6.18e+04	2.91e+04
C(county)[T.166]	-1.346e+04	1.02e+04	-1.321	0.187	-3.34e+04	6513.690
C(county)[T.167]	-3.164e+04	3.52e+04	-0.898	0.369	-1.01e+05	3.74e+04

C(county)[T.168]	1.308e+04	7699.611	1.699	0.089	-2013.719	2.82e+04
C(county)[T.169]	1.863e+04	1.24e+04	1.501	0.134	-5709.855	4.3e+04
C(county)[T.170]	-1.411e+04	1.53e+04	-0.925	0.355	-4.4e+04	1.58e+04
C(county)[T.171]	2.274e+04	7784.588	2.921	0.004	7479.646	3.8e+04
C(county)[T.172]	-2.843e+04	3.52e+04	-0.807	0.420	-9.75e+04	4.06e+04
C(county)[T.173]	-2.254e+04	2.58e+04	-0.872	0.383	-7.32e+04	2.81e+04
C(county)[T.174]	-2.096e+04	1.73e+04	-1.215	0.225	-5.48e+04	1.29e+04
C(county)[T.176]	-1.22e+04	1.92e+04	-0.637	0.524	-4.97e+04	2.53e+04
C(county)[T.177]	-4760.6029	7530.601	-0.632	0.527	-1.95e+04	1e+04
C(county)[T.178]	-1711.2126	1.03e+04	-0.166	0.868	-2.2e+04	1.85e+04
C(county)[T.179]	4219.8694	9883.878	0.427	0.669	-1.52e+04	2.36e+04
C(county)[T.180]	-1.597e+04	7868.741	-2.029	0.042	-3.14e+04	-542.884
C(county)[T.181]	5868.3202	1.26e+04	0.467	0.641	-1.88e+04	3.05e+04
C(county)[T.182]	-1.647e+04	8869.718	-1.856	0.063	-3.39e+04	923.103
C(county)[T.183]	1689.2579	4383.873	0.385	0.700	-6905.058	1.03e+04
C(county)[T.184]	-1.73e+04	4003.987	-4.321	0.000	-2.52e+04	-9453.606
C(county)[T.185]	-1.784e+04	1.63e+04	-1.091	0.275	-4.99e+04	1.42e+04
C(county)[T.186]	-8595.9211	1.16e+04	-0.741	0.458	-3.13e+04	1.41e+04
C(county)[T.187]	-2689.6844	9530.338	-0.282	0.778	-2.14e+04	1.6e+04
C(county)[T.188]	-1.237e+04	3835.130	-3.225	0.001	-1.99e+04	-4848.471
C(county)[T.189]	-1.722e+04	5848.270	-2.944	0.003	-2.87e+04	-5751.363
C(county)[T.190]	-3.427e+04	1.6e+04	-2.137	0.033	-6.57e+04	-2837.788
C(county)[T.191]	-8050.9016	2.08e+04	-0.387	0.699	-4.89e+04	3.28e+04
C(county)[T.192]	7663.4616	1.01e+04	0.761	0.447	-1.21e+04	2.74e+04
C(county)[T.193]	-1361.4111	1.14e+04	-0.120	0.905	-2.36e+04	2.09e+04
C(county)[T.194]	4084.3841	1.47e+04	0.278	0.781	-2.47e+04	3.29e+04
C(county)[T.195]	-1.795e+04	1.04e+04	-1.723	0.085	-3.84e+04	2472.375
C(county)[T.196]	-4086.9677	7126.227	-0.574	0.566	-1.81e+04	9883.567
C(county)[T.197]	9241.7078	6204.356	1.490	0.136	-2921.554	2.14e+04
C(county)[T.198]	1.224e+04	8125.489	1.507	0.132	-3684.922	2.82e+04
C(county)[T.199]	-2.006e+04	9274.975	-2.163	0.031	-3.82e+04	-1877.735
C(county)[T.200]	-1.02e+04	3245.605	-3.143	0.002	-1.66e+04	-3838.033
C(county)[T.201]	-1.703e+04	1.64e+04	-1.038	0.300	-4.92e+04	1.51e+04
C(county)[T.202]	5598.3153	2.45e+04	0.229	0.819	-4.24e+04	5.36e+04
C(county)[T.203]	-6971.9664	1.67e+04	-0.419	0.675	-3.96e+04	2.57e+04
C(county)[T.204]	-1.434e+04	1.07e+04	-1.341	0.180	-3.53e+04	6624.867
C(county)[T.205]	-2.182e+04	6996.633	-3.119	0.002	-3.55e+04	-8104.263
C(county)[T.206]	-3.533e+04	1.64e+04	-2.153	0.031	-6.75e+04	-3161.396
C(county)[T.209]	7696.7394	1.46e+04	0.527	0.599	-2.1e+04	3.64e+04
C(county)[T.210]	-2.404e+04	2.88e+04	-0.835	0.404	-8.05e+04	3.24e+04
C(county)[T.211]	2067.1585	8006.142	0.258	0.796	-1.36e+04	1.78e+04
C(county)[T.213]	-2.644e+04	2.55e+04	-1.035	0.301	-7.65e+04	2.36e+04
C(county)[T.214]	-2.663e+04	1.37e+04	-1.940	0.052	-5.35e+04	274.795
C(county)[T.215]	-1.307e+04	1.21e+04	-1.078	0.281	-3.68e+04	1.07e+04
C(county)[T.216]	-9510.2322	5786.004	-1.644	0.100	-2.09e+04	1832.875
C(county)[T.217]	-2.662e+04	1.04e+04	-2.563	0.010	-4.7e+04	-6257.826
C(county)[T.218]	-2.16e+04	1.81e+04	-1.195	0.232	-5.7e+04	1.38e+04
C(county)[T.219]	-3113.2930	8873.329	-0.351	0.726	-2.05e+04	1.43e+04
C(county)[T.220]	2.191e+04	1.2e+04	1.818	0.069	-1712.077	4.55e+04
C(county)[T.221]	1.036e+04	1.25e+04	0.832	0.406	-1.41e+04	3.48e+04
C(county)[T.222]	-1.297e+04	1.11e+04	-1.169	0.243	-3.47e+04	8789.494
C(county)[T.223]	-1.721e+04	5639.229	-3.053	0.002	-2.83e+04	-6158.489
C(county)[T.224]	-1.102e+04	1.57e+04	-0.701	0.483	-4.18e+04	1.98e+04
C(county)[T.225]	-347.2929	2.15e+04	-0.016	0.987	-4.25e+04	4.18e+04
C(county)[T.226]	-6743.7161	7207.321	-0.936	0.349	-2.09e+04	7385.797
C(county)[T.227]	-1.871e+04	1.18e+04	-1.584	0.113	-4.19e+04	4442.638
C(county)[T.228]	3.605e+04	1.85e+04	1.943	0.052	-317.169	7.24e+04
C(county)[T.229]	-1.079e+04	2.07e+04	-0.521	0.603	-5.14e+04	2.98e+04

C(county)[T.230]	-1.919e+04	6749.998	-2.843	0.004	-3.24e+04	-5959.853
C(county)[T.231]	-7785.4723	7866.613	-0.990	0.322	-2.32e+04	7636.543
C(county)[T.232]	-2.419e+04	9240.849	-2.617	0.009	-4.23e+04	-6069.259
C(county)[T.233]	-2.042e+04	9194.857	-2.221	0.026	-3.84e+04	-2397.294
C(county)[T.234]	-2.54e+04	3.37e+04	-0.753	0.451	-9.15e+04	4.07e+04
C(county)[T.235]	-2.558e+04	3.43e+04	-0.745	0.457	-9.29e+04	4.18e+04
C(county)[T.236]	-1.215e+04	9066.209	-1.340	0.180	-2.99e+04	5620.742
C(county)[T.237]	-1.191e+04	4869.450	-2.445	0.015	-2.15e+04	-2359.185
C(county)[T.238]	-1.574e+04	7743.041	-2.032	0.042	-3.09e+04	-556.010
C(county)[T.239]	-2.298e+04	2.33e+04	-0.987	0.324	-6.86e+04	2.27e+04
C(county)[T.240]	5577.5142	9917.460	0.562	0.574	-1.39e+04	2.5e+04
C(county)[T.241]	-8698.4006	6719.914	-1.294	0.196	-2.19e+04	4475.582
C(county)[T.242]	-1.48e+04	1.38e+04	-1.075	0.283	-4.18e+04	1.22e+04
C(county)[T.243]	-2.601e+04	1.51e+04	-1.722	0.085	-5.56e+04	3606.381
C(county)[T.244]	-1.79e+04	1.55e+04	-1.153	0.249	-4.83e+04	1.25e+04
C(county)[T.245]	-1.127e+04	1.47e+04	-0.767	0.443	-4.01e+04	1.76e+04
C(county)[T.246]	1396.0584	4336.583	0.322	0.748	-7105.549	9897.666
C(county)[T.247]	-7809.7675	6125.901	-1.275	0.202	-1.98e+04	4199.688
C(county)[T.248]	-4669.3444	9431.966	-0.495	0.621	-2.32e+04	1.38e+04
C(county)[T.249]	-1.1e+04	3.64e+04	-0.302	0.763	-8.25e+04	6.05e+04
C(county)[T.250]	2.682e+04	1.14e+04	2.356	0.018	4507.158	4.91e+04
C(county)[T.251]	-1.772e+04	1.04e+04	-1.708	0.088	-3.81e+04	2615.894
C(county)[T.252]	-2.187e+04	2.18e+04	-1.003	0.316	-6.46e+04	2.09e+04
C(county)[T.253]	-1.797e+04	1.71e+04	-1.050	0.294	-5.15e+04	1.56e+04
C(county)[T.254]	-1.629e+04	1.34e+04	-1.214	0.225	-4.26e+04	1e+04
C(county)[T.256]	-4328.4864	1.43e+04	-0.303	0.762	-3.23e+04	2.36e+04
C(county)[T.257]	-1.819e+04	5031.865	-3.615	0.000	-2.81e+04	-8323.053
C(county)[T.258]	-2.719e+04	2.02e+04	-1.346	0.178	-6.68e+04	1.24e+04
C(county)[T.259]	7.088e+04	1.94e+04	3.660	0.000	3.29e+04	1.09e+05
C(county)[T.260]	-2.466e+04	1.54e+04	-1.599	0.110	-5.49e+04	5578.948
C(county)[T.261]	-1.29e+04	1.03e+04	-1.250	0.212	-3.31e+04	7340.884
C(county)[T.262]	-2.291e+04	3.43e+04	-0.667	0.505	-9.02e+04	4.44e+04
C(county)[T.263]	9962.7019	1.36e+04	0.732	0.464	-1.67e+04	3.67e+04
C(county)[T.264]	-6547.2832	1.28e+04	-0.510	0.610	-3.17e+04	1.86e+04
C(county)[T.265]	-1.97e+04	1.14e+04	-1.733	0.083	-4.2e+04	2591.085
C(county)[T.266]	9962.7019	3.64e+04	0.273	0.785	-6.15e+04	8.14e+04
C(county)[T.267]	-1.05e+04	2.54e+04	-0.414	0.679	-6.02e+04	3.92e+04
C(county)[T.268]	-8977.5514	1.64e+04	-0.547	0.584	-4.11e+04	2.32e+04
C(county)[T.269]	-2.085e+04	8575.051	-2.432	0.015	-3.77e+04	-4040.781
C(county)[T.270]	-1.725e+04	1.87e+04	-0.921	0.357	-5.4e+04	1.95e+04
C(county)[T.271]	-1.085e+04	1.68e+04	-0.645	0.519	-4.39e+04	2.22e+04
C(county)[T.272]	-1.703e+04	9648.288	-1.765	0.078	-3.59e+04	1888.592
C(county)[T.273]	-3.333e+04	2.85e+04	-1.168	0.243	-8.93e+04	2.26e+04
C(county)[T.274]	-1.11e+04	5089.398	-2.181	0.029	-2.11e+04	-1123.537
C(county)[T.275]	-4769.1776	1.25e+04	-0.381	0.703	-2.93e+04	1.97e+04
C(county)[T.276]	-1.814e+04	1.81e+04	-1.004	0.315	-5.36e+04	1.73e+04
C(county)[T.277]	-2.956e+04	3.11e+04	-0.950	0.342	-9.05e+04	3.14e+04
C(county)[T.278]	-9478.9414	2.06e+04	-0.459	0.646	-4.99e+04	3.1e+04
C(county)[T.279]	-1.548e+04	1.99e+04	-0.776	0.438	-5.46e+04	2.36e+04
C(county)[T.280]	-1.343e+04	6815.226	-1.970	0.049	-2.68e+04	-67.933
C(county)[T.281]	-1.266e+04	3.24e+04	-0.390	0.696	-7.63e+04	5.09e+04
C(county)[T.282]	-5823.9351	1.3e+04	-0.447	0.655	-3.14e+04	1.97e+04
C(county)[T.283]	-1.498e+04	1.08e+04	-1.389	0.165	-3.61e+04	6159.586
C(county)[T.284]	-1.731e+04	8871.403	-1.951	0.051	-3.47e+04	79.927
C(county)[T.285]	-2.878e+04	2.31e+04	-1.247	0.212	-7.4e+04	1.64e+04
C(county)[T.286]	7886.7762	1.26e+04	0.628	0.530	-1.68e+04	3.25e+04
C(county)[T.287]	-1.828e+04	5568.186	-3.282	0.001	-2.92e+04	-7361.137
C(county)[T.288]	6217.2362	1.92e+04	0.324	0.746	-3.14e+04	4.38e+04

C(county)[T.289]	-1.065e+04	1.2e+04	-0.886	0.376	-3.42e+04	1.29e+04
C(county)[T.290]	-1.516e+04	7708.924	-1.966	0.049	-3.03e+04	-42.553
C(county)[T.291]	-9430.5569	1.47e+04	-0.643	0.521	-3.82e+04	1.93e+04
C(county)[T.292]	-2.032e+04	7927.306	-2.563	0.010	-3.59e+04	-4774.946
C(county)[T.293]	-1.053e+04	1.46e+04	-0.722	0.470	-3.91e+04	1.81e+04
C(county)[T.294]	3.081e+04	1.07e+04	2.886	0.004	9877.836	5.17e+04
C(county)[T.295]	-5745.5790	1.91e+04	-0.301	0.763	-4.32e+04	3.17e+04
C(county)[T.296]	-2113.6831	1.96e+04	-0.108	0.914	-4.05e+04	3.62e+04
C(county)[T.297]	-1.294e+04	4780.000	-2.707	0.007	-2.23e+04	-3570.673
C(county)[T.298]	-1.254e+04	3.12e+04	-0.402	0.687	-7.37e+04	4.86e+04
C(county)[T.299]	-2.165e+04	3.24e+04	-0.667	0.505	-8.53e+04	4.19e+04
C(county)[T.300]	-1.465e+04	1.13e+04	-1.297	0.195	-3.68e+04	7495.261
C(county)[T.301]	-3.546e+04	3.52e+04	-1.007	0.314	-1.05e+05	3.36e+04
C(county)[T.302]	-561.3863	8241.124	-0.068	0.946	-1.67e+04	1.56e+04
C(county)[T.303]	-2.213e+04	8027.456	-2.757	0.006	-3.79e+04	-6397.232
C(county)[T.304]	-1.696e+04	9517.227	-1.782	0.075	-3.56e+04	1695.665
C(county)[T.305]	-1941.4937	7830.771	-0.248	0.804	-1.73e+04	1.34e+04
C(county)[T.306]	-3.482e+04	2.16e+04	-1.611	0.107	-7.72e+04	7552.428
C(county)[T.307]	-1.323e+04	1.06e+04	-1.245	0.213	-3.41e+04	7603.790
C(county)[T.308]	9.756e+04	1.8e+04	5.417	0.000	6.22e+04	1.33e+05
C(county)[T.309]	-1.72e+04	5528.470	-3.112	0.002	-2.8e+04	-6364.578
C(county)[T.310]	-8923.6981	3.64e+04	-0.245	0.807	-8.04e+04	6.25e+04
C(county)[T.311]	-9085.9943	1.09e+04	-0.831	0.406	-3.05e+04	1.24e+04
C(county)[T.312]	-4418.6964	1.07e+04	-0.412	0.680	-2.54e+04	1.66e+04
C(county)[T.313]	-2.249e+04	4.66e+04	-0.483	0.629	-1.14e+05	6.88e+04
C(county)[T.314]	-1.72e+04	7780.475	-2.211	0.027	-3.25e+04	-1947.024
C(county)[T.315]	1057.8405	1.13e+04	0.093	0.926	-2.11e+04	2.32e+04
C(county)[T.316]	-3703.2330	7330.899	-0.505	0.613	-1.81e+04	1.07e+04
C(county)[T.317]	1.584e+04	1.01e+04	1.563	0.118	-4032.805	3.57e+04
C(county)[T.318]	-6428.9638	9307.531	-0.691	0.490	-2.47e+04	1.18e+04
C(county)[T.319]	4465.5629	1.54e+04	0.289	0.772	-2.58e+04	3.47e+04
C(county)[T.320]	-1.247e+04	1.64e+04	-0.762	0.446	-4.45e+04	1.96e+04
C(county)[T.321]	-1.738e+04	7553.474	-2.301	0.021	-3.22e+04	-2573.721
C(county)[T.322]	-3579.7346	7080.094	-0.506	0.613	-1.75e+04	1.03e+04
C(county)[T.323]	-1.61e+04	6210.453	-2.593	0.010	-2.83e+04	-3929.679
C(county)[T.324]	1339.5164	8530.644	0.157	0.875	-1.54e+04	1.81e+04
C(county)[T.325]	-2.022e+04	2.43e+04	-0.832	0.405	-6.79e+04	2.74e+04

Omnibus:	2406.664	**Durbin-Watson:**	2.009
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	18545.732
Skew:	2.003	**Prob(JB):**	0.00
Kurtosis:	11.232	**Cond. No.**	156.

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Coefficients of simplified race are near to 0, and the p-values are 1, indicating it is **not statistically significant**, i.e. **simplified race is balanced**.

Exercise 13： regress annual earnings on just having a college eduction

In [ ]:
```
smf.wls(
    "annual_earnings ~ has_college", matched_data, weights=matched_data["weights"]
).fit().summary()
```

Out[ ]:

WLS Regression Results

Dep. Variable:	annual_earnings	**R-squared:**	0.058
Model:	WLS	**Adj. R-squared:**	0.057
Method:	Least Squares	**F-statistic:**	324.1
Date:	Wed, 01 Mar 2023	**Prob (F-statistic):**	2.19e-70
Time:	22:48:45	**Log-Likelihood:**	-61753.
No. Observations:	5310	**AIC:**	1.235e+05
Df Residuals:	5308	**BIC:**	1.235e+05
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	3.909e+04	351.293	111.287	0.000	3.84e+04	3.98e+04
has_college	1.374e+04	763.203	18.003	0.000	1.22e+04	1.52e+04

Omnibus:	2934.035	**Durbin-Watson:**	2.006
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	33100.529
Skew:	2.424	**Prob(JB):**	0.00
Kurtosis:	14.230	**Cond. No.**	2.58

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

- The annual earnings of people who has a college degree is $13740 higher than those who don't.
- It is lower than the initial estimate using the raw CPS data ($14158.50).

Exercise 14: include other matching variables as controls

In [ ]:
```
matched_data.columns
```

Out[ ]:
```
Index(['annual_earnings', 'female', 'simplified_race', 'has_college', 'county',
       'class94', 'age_group', 'match_group', 'match_group_size', 'weights'],
      dtype='object')
```

In [ ]:
```
# we dropped county
smf.wls(
    "annual_earnings ~ has_college + female + C(simplified_race) + C(class94) + age_group", \
        matched_data, weights=matched_data["weights"]
).fit().summary()
```

Out[ ]:

WLS Regression Results

Dep. Variable:	annual_earnings	**R-squared:**	0.165
Model:	WLS	**Adj. R-squared:**	0.162
Method:	Least Squares	**F-statistic:**	74.51
Date:	Wed, 01 Mar 2023	**Prob (F-statistic):**	1.64e-194
Time:	22:48:45	**Log-Likelihood:**	-61433.
No. Observations:	5310	**AIC:**	1.229e+05
Df Residuals:	5295	**BIC:**	1.230e+05
Df Model:	14		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	5.179e+04	1748.995	29.611	0.000	4.84e+04	5.52e+04
C(simplified_race)[T.1.0]	-7900.4191	957.233	-8.253	0.000	-9776.990	-6023.848
C(simplified_race)[T.2.0]	-3930.8604	995.077	-3.950	0.000	-5881.621	-1980.100
C(simplified_race)[T.3.0]	-1454.4252	1072.357	-1.356	0.175	-3556.688	647.837
C(class94)[T.1]	-1.355e+04	1867.099	-7.260	0.000	-1.72e+04	-9894.581
C(class94)[T.2]	-1.437e+04	1922.307	-7.474	0.000	-1.81e+04	-1.06e+04
C(class94)[T.3]	-1.542e+04	1617.269	-9.536	0.000	-1.86e+04	-1.23e+04
C(class94)[T.4]	-1.304e+04	1840.039	-7.085	0.000	-1.66e+04	-9429.670
age_group[T.1.0]	8466.8167	846.944	9.997	0.000	6806.457	1.01e+04
age_group[T.2.0]	1.112e+04	901.314	12.341	0.000	9356.585	1.29e+04
age_group[T.3.0]	1.148e+04	937.844	12.236	0.000	9637.084	1.33e+04
age_group[T.4.0]	8989.7474	1159.161	7.755	0.000	6717.315	1.13e+04
age_group[T.5.0]	1.292e+04	2920.324	4.423	0.000	7190.157	1.86e+04
has_college	1.374e+04	719.437	19.099	0.000	1.23e+04	1.52e+04
female	-8574.8140	602.487	-14.232	0.000	-9755.937	-7393.691

Omnibus:	2962.601	**Durbin-Watson:**	1.992
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	41404.049
Skew:	2.373	**Prob(JB):**	0.00
Kurtosis:	15.830	**Cond. No.**	18.2

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Since we dropped county during matching, we added all other variables except county. The coefficient of has_college doesn't change.

Exercise 15: exact matches

In [ ]:
```
model3 = dame_flame.matching.DAME(repeats=False, verbose=3, want_pe=True, early_stop_iterations=0)
model3.fit(
    cps_2,
    treatment_column_name="has_college",
    outcome_column_name="annual_earnings",
)
result3 = model3.predict(cps_2)
```

```
Completed iteration 0 of matching
	Number of matched groups formed in total:  370
	Unmatched treated units:  644 out of a total of  1150 treated units
	Unmatched control units:  3187 out of a total of  4365 control units
	Number of matches made this iteration:  1684
	Number of matches made so far:  1684
	Covariates dropped so far:  set()
	Predictive error of covariate set used to match:  1199312680.0957854
1684 units matched. We stopped after iteration 0
```

In [ ]:
```
matched_data_2 = get_dataframe(model3, result3)
matched_data_2.head()
```

Out[ ]:

	annual_earnings	female	simplified_race	has_college	county	class94	age_group	match_group	match_group_size	weights
0	42900.0	1	0.0	0	10	3	3.0	59.0	5.0	0.582016
2	20020.0	0	0.0	1	8	3	4.0	52.0	3.0	1.000000
4	73860.8	0	0.0	1	24	3	1.0	106.0	7.0	1.000000
6	32760.0	1	0.0	0	90	3	0.0	276.0	3.0	1.164032
12	27040.0	0	0.0	0	10	3	0.0	54.0	8.0	0.332580

Exercise 16: regress annual earnings on just having a college eduction with new matched data

In [ ]:
```python
smf.wls(
    "annual_earnings ~ has_college", matched_data_2, weights=matched_data_2["weights"]
).fit().summary()
```

Out[ ]:

WLS Regression Results

Dep. Variable:	annual_earnings	**R-squared:**	0.049
Model:	WLS	**Adj. R-squared:**	0.048
Method:	Least Squares	**F-statistic:**	86.65
Date:	Wed, 01 Mar 2023	**Prob (F-statistic):**	3.92e-20
Time:	22:48:46	**Log-Likelihood:**	-19512.
No. Observations:	1684	**AIC:**	3.903e+04
Df Residuals:	1682	**BIC:**	3.904e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
Intercept	3.914e+04	664.386	58.907	0.000	3.78e+04	4.04e+04
has_college	1.128e+04	1212.039	9.308	0.000	8904.805	1.37e+04

Omnibus:	855.250	**Durbin-Watson:**	2.037
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	6653.000
Skew:	2.256	**Prob(JB):**	0.00
Kurtosis:	11.629	**Cond. No.**	2.42

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

With new matched data, the coefficient of has_college is $11280, meaningholdingothervariablesconstant, theannualearningsofpeoplewhohasacollegedegreeis$11280 higher than those who don't. The value is lower than what we got from Exercise 13 ($13740), suggesting that the salary gap between people with and without college degree is smaller than when we restrict matching to exact matches.