Python Data Analysis and Visualization

Claudia Carroll, TRIADS

**Homework Exercises: Data Aggregation**

1. Write the code to find the mean value of a group of three columns ‘Travel’, ‘Groceries’ and ‘Coffee’ in a dataframe df.

df.groupby([‘Travel’, ‘Groceries’, ‘Coffee’]).mean()

1. Write the code to merge the columns in a dataframe df2 that have the same column headers as those in a dataframe df to df and save it to a third dataframe df 3. Set the merge to happen on a column titled ‘degree’.

df3 = df.merge(df2, how=‘left’, on=‘degree’)

1. Can you do the same thing as question 4, but using concatenate? You may need a few extra steps!

common\_degrees = df2[df2['degree'].isin(df['degree'])].index

df2\_subset = df2.loc[common\_degrees]

df3 = pd.concat([df.set\_index('degree'),df2\_subset.set\_index('degree')], axis=1)

df2\_subset.set\_index('degree')], axis=1)