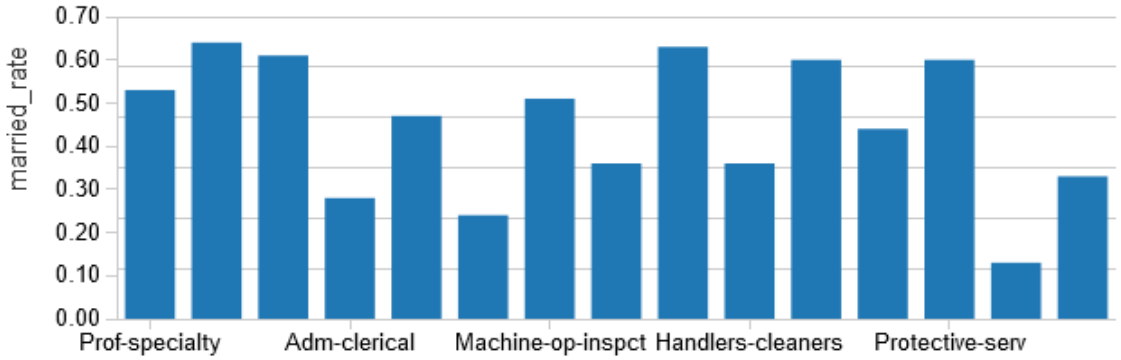
I performed data analysis work on the U.S. census dataset concerning adults in the U.S., but more specifically their ages, workclasses, educations, marital statuses, occupations, relationships, races, sexes, capital gain, capital loss, number of hours worked per week, native countries and incomes. CMD4 creates a dataset holding all this information, and CMD5 returns the first 100 results in this dataset (unordered). CMD6 returned a bar graph representing the marital status rate of every occupation (i.e., the married rate, widow rate, divorce rate, separated rate and bachelor rate), with the largest occupation listed first – the bar graph looked like this:



CMD7 returned the education level with the highest bachelor rate, which ended up being a 12th grade education with a bachelor rate of 0.54 (54%). CMD8 then created a dataframe with the dataset and CMD9 displayed the schema for it (i.e., what values could be contained in each column and whether no value was acceptable or not). CMD10 listed the top 5 occupations with the highest divorce rates, which were as follows:

occupation	divorced_rate
Adm-clerical	0.2172990183072433
Priv-house-serv	0.18791946308724833
Other-service	0.15204855842185128
Tech-support	0.15086206896551724
Exec-managerial	0.1485489424495819

CMD11 was the last command, and it listed the top 5 education levels with the highest amount of people with incomes over $50,000, which were as follows:

education	count
Bachelors	2221
HS-grad	1675
Some-college	1387
Masters	959
Prof-school	423