MSSP 607: Homework 1

Part 1

In no less than 100 words, why might someone want to look at this data? Why would it matter?

Data on school district finance are of significance in educational funding analyses. From the microscope, this data gives information on school district's funding situation, which can be used to learn the specifics about school districts' revenues and budgets. This data also reflects states' features of demographics and administrative division. From the macro perspectives, this data is useful for data scientists to evaluate the school funding distribution across the country and to answer questions such as whether the federal funding is distributed evenly in proportion to a state's population, whether a state is overfunded or underfunded by calculating the budget deficits and whether changes in school finance policies are required to better support schools and education. Together with other datasets, such as states' GDP, taxation, local living prices, quantitative representation of education quality, and so on, this data can help to analyze the correlations between school district finance and economy, between school funding and education quality, etc. Comparing to federal expenses on education, this data can also examine whether the federal budget plan were well implemented. Analyses based on this data can help to navigate adjustments to future policy making and budget plans.

Part 2 (Q1 - Q4)

1. In our dataset, what percentage of school districts nationwide received more than $10,000,000 in local funding?

30.3%

About 30.3% of school districts nationwide received more than $10,000,000 from local revenue in 2014.

2. In our dataset, among school districts with a total budget (expenses) of at least $1 billion, which has the shortest name? What state is that district in?

Dade in **Florida**

Among school districts with expenses at least $1 billion in 2014, Dade district has the shortest name, and it is in the state of Florida.

3. Hawaii has only one school district, which covers the entire state. Across the rest of the United States, how many school districts had greater expenses than Hawaii did statewide?

8

There are 8 school districts that had greater expenses than Hawaii did statewide in 2014.

4. The budget deficit of a school district is equal to its expenses minus its revenue. Which school district in Pennsylvania had the largest budget deficit? And how large was it?

PENN HILLS SD: **32789000**
In 2014, Penn Hills SD had the largest budget deficit in Pennsylvania and its deficit was 32,789,000 (32,789 thousand) USD.

5. **#[Only for practice & fun!]** New England comprises the states Maine, Vermont, New Hampshire, Rhode Island, Massachusetts, and Connecticut. To calculate per capita statistics, you divide the total amount by the total population size. Which state in New England gave the most non-federal money (state and local combined) to schools? How much did they give per capita?
Massachusetts: **2498.29** USD per capita
In New England, Massachusetts gave the most non-federal money to schools (which was 16,357,898,000 USD) and they gave 2498.29 USD per capita in 2014.

6. **#[for practice]** The mean size of a school district in a state is equal to the total population divided by the total number of school districts. Which state had the smallest mean school district size? How many people, on average, live in each school district in that state?
Vermont: **1955** people each district
Vermont, with 320 school districts, had the smallest mean school district size of 1955 people living in each school district in 2014.

Part 3
Which state receives the "best deal" in school funding from the federal government? Why? Provide the source code you wrote to generate any statistics supporting your argument, and provide your answer in a professional and appropriate format that backs up your reasoning.

I think Louisiana received the "best deal" in school funding from the federal government in 2014.

To evaluate a "deal" in federal school funding among states, factors such as the total amount of federal school funding, states' federal funding per school district, states' federal school funding per capita, and the proportion federal funds took of states' total revenues should be taken into consideration.

1. Amount of federal school funding:

```
def p3_federal_state_sum(districts_list, state_populations):
    state_federal = {}
    for k in state_populations:
        state_federal_sum = 0
        for district_dict in districts_list:
            if district_dict['state'] == k:
                state_federal_sum  += district_dict['federal_revenue']
        state_federal[k] = state_federal_sum
    return state_federal


def p3_top5_federal_funding(state_federal_sum):
    state_fed_sum_items = list(state_federal_sum.items())
    sort_fed_sum = sorted(state_fed_sum_items, key = lambda item:item[1], reverse=True )
    print(f"P3#1 top 5 most federal funded states (k): \n{sort_fed_sum[0:5]}")
    print("-"*100)
    return sort_fed_sum


p3_top5_federal_funding(p3_federal_state_sum(districts_list, state_populations))
```

For the first part, I calculated the 5 states that have the highest amounts of federal funding and they are California, Texas, New York, Florida, and Illinois. Due to the fact that states are of different economic volumes, this result cannot directly contribute to the final evaluation of a federal funding deal and, thus, it is only for reference.

2. Federal funding per school district:

One way to evaluate federal funding deal among states of different demographic features is to calculate how much money of the federal funding each school district shares, by dividing the sum of federal revenue by the total number of school district in each state. The result of calculation of the top 5 states of highest federal funding per school district is Hawaii (286988.0 k/sd), Florida (46448.16 k/sd), Maryland (34001.38 k/sd), Nevada (20288.26 k/sd), and Louisiana (10766.96 k/sd). One tricky thing about this result is that it is closely related to a state's administrative division. For example, there is only one school district in entire Hawaii State, which naturally leads to a high per school district federal revenue. And that is why federal funding per school district cannot serve as the only factor to decide which state has the "best deal".

```
def p3_state_sd_count(districts_list, state_populations):
    state_sd_count = {}
    for k in state_populations:
        sd_count = 0
        for district_dict in districts_list:
            if district_dict['state'] == k:
                sd_count += 1
        if sd_count > 0:
            state_sd_count[k] = sd_count
    return state_sd_count


def p3_top5_federal_per_sd(state_sum,sd_count):
    federal_sd = {}
    for k1 in state_sum:
        for k2 in sd_count:
            if k1 == k2:
                federal_per_sd = round(state_sum [k1]/sd_count[k2],2)
        federal_sd[k1] = federal_per_sd
    federal_sd_items = list(federal_sd.items())
    sort_fed_sd = sorted(federal_sd_items, key= lambda item: item[1], reverse = True)
    print(f"P3#2 top 5 states of highest federal funding per district (k):
\n{sort_fed_sd[0:5]}")
    print("-" * 100)
    return sort_fed_sd[0:5]


p3_top5_federal_per_sd(p3_federal_state_sum(districts_list, state_populations),
                       p3_state_sd_count(districts_list, state_populations))
```

3. Federal funding per capita:

Another demographic feature that affects a state's federal funding in schools is the state's population, as school funding should be in proportion to the population of the

state. Dividing the amount of federal funding by the entire population results in the federal funding the government gives per capita. The top 5 states that had highest federal revenue per capita were Alaska (440.0 dollar/person), District of Columbia (340.0 dollar/person), Louisiana (300.0 dollar/person), Texas (230.0 dollar/person), and New Mexico (230.0 dollar/person). So for these 5 states, each person has more access to federal funds than the residents from other states and households have more expensive schooling resources comparing to households in other state with same number of school-age children.

```
def p3_top5_fed_per_capita(state_sum,state_population):
    fed_capita = {}
    for k1 in state_sum:
        for k2 in state_population:
            if k1 == k2:
                fed_per_capita = round(state_sum[k1]/state_population[k2],2)*1000
        fed_capita[k1] = fed_per_capita
    fed_capita_items = list(fed_capita.items())
    sort_fed_cap = sorted(fed_capita_items, key = lambda item: item[1], reverse = True)
    print(f"P3#3 top 5 states of highest federal funding per capita (k): \n{sort_fed_cap[0:5]}")
    print("-" * 100)
    return sort_fed_cap[0:5]


p3_top5_fed_per_capita(p3_federal_state_sum(districts_list, state_populations),
state_populations)
```

4. Federal funding proportion:

```
def p3_state_revenue_sum(districts_list, state_populations):
    state_revenue = {}
    for k in state_populations:
        state_revenue_sum = 0
        for district_dict in districts_list:
            if district_dict['state'] == k:
                state_revenue_sum  += district_dict['revenue']
        state_revenue[k] = state_revenue_sum
    return state_revenue


def p3_top5_fed_proportion(state_fed_sum,state_revenue_sum):
    fed_prop = {}
    for k1 in state_revenue_sum:
        for k2 in state_fed_sum:
            if k1 == k2:
                if state_revenue_sum[k1] != 0:
                    fed_proportion = round((state_fed_sum[k2]/state_revenue_sum[k1])*100,2)
        fed_prop[k1] = fed_proportion
    fed_prop_items = list(fed_prop.items())
    sort_fed_prop = sorted(fed_prop_items, key = lambda item:item[1], reverse=True)
    print(f"P3#4 top 5 state of highest federal funding proportion (%): \n{sort_fed_prop[0:5]}")
    print("-" * 100)
    return sort_fed_prop[0:5]


p3_top5_fed_proportion(p3_federal_state_sum(districts_list, state_populations),
                       p3_state_revenue_sum(districts_list, state_populations))
```

Of what percentage a state's total revenue is federal funding is also an indicator to evaluate the federal funding deal. The higher the percentage the federal revenue takes, the greater the role that the government played in funding schools. The 5 states that are most heavily federal funded in 2014 were Louisiana (15.18%), Mississippi (14.84%), South Dakota (13.84%), New Mexico (12.72%), and Arizona (12.38%). While the proportion federal funding takes is also relevant to the amount of state and local revenue, thus to state and local economy, it is undeniable that the states of higher proportion benefit more from federal funds.

5. Best deal

With all of the factors (2-4) above considered, Louisiana was the only state that ranked top 5 in federal funding per school district, federal funding per capita and the percentage of federal funding in total revenue. Thus, I think Louisiana was the state that had the "best deal" in school funding from the federal government in 2014.

```
def p3_best_deal(t5_fed_per_sd,t5_fed_per_cap,t5_fed_ratio):
    best_deal = []
    for item1 in t5_fed_per_sd:
        for item2 in t5_fed_per_cap:
            if item1[0] == item2[0]:
                best_deal_candidate = item1[0]
                for item3 in t5_fed_ratio:
                    if best_deal_candidate == item3[0]:
                        best_deal.append(best_deal_candidate)
    print(f"P3#5 best deal candidate(s): \n{best_deal}")
    return best_deal


p3_best_deal(p3_top5_federal_per_sd(p3_federal_state_sum(districts_list, state_populations),
                                    p3_state_sd_count(districts_list, state_populations)),
             p3_top5_fed_per_capita(p3_federal_state_sum(districts_list,
state_populations),state_populations),
             p3_top5_fed_proportion(p3_federal_state_sum(districts_list, state_populations),
                                    p3_state_revenue_sum(districts_list, state_populations)))
```

Extra Credit

<u>In no less than 300 words, describe some other questions you think could be interesting when researching school districts, their sources of funding, and their expenses. Why are those questions interesting? Are the values in the homework file enough to answer those questions? If not, what additional columns would you need to collect in order to answer them?</u>

Question 1: How is a state's economy reflected in its school distribution funding?

This question is interesting because I have an assumption that the better being a state's economy is, the lower the percentage of total revenue amount the federal funding takes. Finding evidence for this assumption may potentially lead to understanding the role federal funding plays in achieving equality in school education across the country. To answer this question, further data will be needed on the state's economy, such as state GDP and GDP per capita, and the states' pricing level so as to differentiate quantity of educational resources and the cost of those resources. Ideally,

in my perception, the government should subsidize the states that are underfunded state-wise and locally more in order to make sure that people have access to as about equal educational resources in general as possible. To be specific, what I would do with the data on state's economic indicators is to divide them by the state's federal funding percentage to see if they are negatively correlated.

Question 2: If a district's school funding per capita is correlated with its quality of school education?

This question is interesting because it finds out how money affects education quality. Logically, the more a state spends on education per capita, the better the education quality (evaluated by education results, student performance and people's degree of satisfaction) should be. If not the case, there could be a problem in the state's educational system, educational policy, or usage of the funds (such as leaky budget effect, and cost on bureaucracy). To answer this question, data to quantify education quality are needed, such as the states' average standard test scores, high school graduate rates, high school graduates' college admission rates and survey results on people's degree of satisfaction. Answering this question is essential to coming up with strategies to improve education quality and solving existing problems. To do so, I would compare a school district's funding to its education quality and students' performance to see if it abides the rule of a positive correlation.