Charter School Budget Analysis

October 12, 2023

1 Budget Analysis

1.1 Abstract

A small charter school network is performing a quantitative budget analysis to calculate the number of filled and vacant positions for all part-time, full-time, and contracted employees as well as deficits and surpluses in department and school budgets to work collaboratively with the Talent department ahead of the upcoming recruiting season to make effective hires.

1.2 Objective

The analysis seeks to answer the following questions:

1. What is the typical length of employment and how does it vary across different departments, campuses, or roles?
2. What is the difference in the projected and actual salaries for both departments and campuses?
3. How many positions were unbudgeted and what is the amount of unbudgeted salaries?

```
[202]:  # importing packages
        import pandas as pd
        import numpy as np
        import datetime as dt
        from dateutil.relativedelta import relativedelta
        import matplotlib.pyplot as plt
```

```
[282]:  # importing budget data
        df = pd.read_csv(r"/Users/scipio/Downloads/Charter_School_Budget_Data.csv")
        df.head()
```

```
[282]:           Name                                   Position       Department  \
        0       Filled                          Safety Specialist        Custodial
        1       Filled                          Safety Specialist        Custodial
        2       Filled                          Safety Specialist        Custodial
        3       VACANT                   Chief Advancement Officer     School Admin
        4   ELIMINATED   School Business Administrator (Part-Time)  Business Office

           Campus  Finance_Budget_Salary  Actual_Budget_Salary  Budgeted  \
        0      HO                47000.0               79577.0  Budgeted
        1      HO                47000.0               79577.0  Budgeted
```

```
2      HO              47000.0            79577.0  Budgeted
3      HO             172500.0                0.0  Budgeted
4      HO              51741.0                0.0  Budgeted

   Employment_Start_Date
0                    NaN
1                    NaN
2                    NaN
3                    NaN
4                    NaN
```

1.3 Analysis

1. What is the typical length of employment and how does it vary across different departments, campuses, or roles?

```
[349]: #Creating datetime object
       current_date = dt.datetime(2023,9,23)

       #subsetting for filled roles based on the 'Employment_Start_Date' column
       employment_length = df[df['Employment_Start_Date'].notnull()]

       #changing data type of 'Employment_Start_Date' column to datetime
       employment_length['Employment_Start_Date'] = pd.
        ↪to_datetime(employment_length['Employment_Start_Date'])

       #inserting the variable 'current_date' into the dataframe
       employment_length['Current_Date'] = current_date

       #Calculating Difference in Years
       employment_length['Employment_Length_Year'] = employment_length.apply(lambda
        ↪row: relativedelta(row['Current_Date'], row['Employment_Start_Date']).years,
        ↪axis=1)

       #Calculating Difference in Months as a decimal
       employment_length['Employment_Length_Month'] = round(employment_length.
        ↪apply(lambda row: relativedelta(row['Current_Date'],
        ↪row['Employment_Start_Date']).months, axis=1)/12,2)

       #Creating a new column with the year and decimal portion
       employment_length['Employment_Length_Overall'] =
        ↪employment_length['Employment_Length_Year'] +
        ↪employment_length['Employment_Length_Month']
```

```
[149]: # Employment Length Statisticla Metrics
       round(employment_length['Employment_Length_Overall'].describe(),3)
```

```
[149]:  count    152.000
        mean       3.274
        std        3.680
        min        0.000
        25%        0.730
        50%        1.830
        75%        4.185
        max       14.000
        Name: Employment_Length_Overall, dtype: float64
```

```
[347]:  # Calculating the percentage of filled positions
        round(df[df['Employment_Start_Date'].notnull()].shape[0]/df.shape[0],2)
```

```
[347]:  0.79
```

79% of the positions are filled in the small charter school network.

```
[150]:  # Visualizing distribution of 'Employment_Length_Years'
        fig, axs = plt.subplots(nrows = 2, ncols = 1, figsize = (5,5))

        employment_length['Employment_Length_Overall'].plot(kind = 'hist', title =␣
         ↪'Employment Length in Years Histogram', ax = axs[0])
        employment_length['Employment_Length_Overall'].plot(kind = 'box', vert = False)
```

```
[150]:  <Axes: >
```
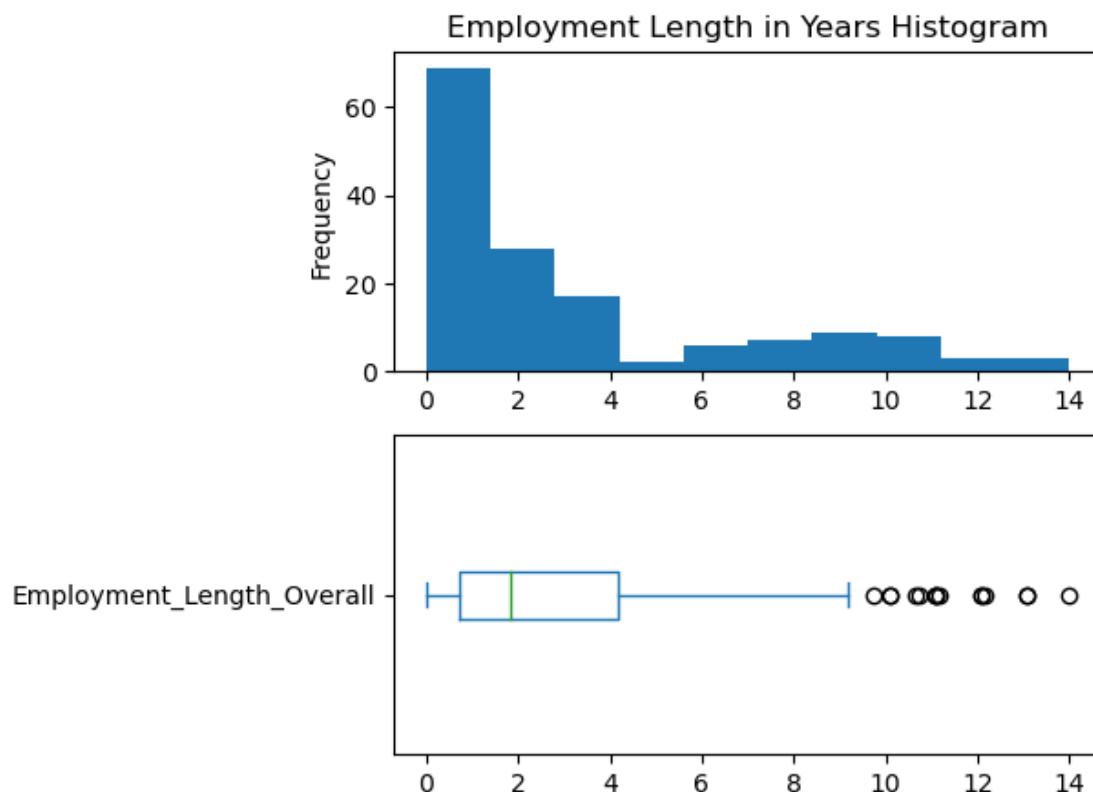
There is a right skew in the histogram above which displays the distribution of employment length in years of employees in the small charter school network. Additionally, based on the boxplot above employees that have been employed in the small charter school network for more than approximately 9 years are outliers. The median length of employment would be a better indicator of the typical length of employment in the charter school network rather than the mean. The typical length of employment at the small charter school network is 1.83 years, approximately 1 year and 10 months.

```
[151]: # Percentage of employees in their first year

       #First year employees
       first_year_employees =␣
        ↪employment_length[employment_length['Employment_Length_Overall'] <1.00] .
        ↪shape[0]

       #total employees
       total_employees = employment_length['Name'].count()

       #percentage of employees in their first year -- first_year_employees/
        ↪total_employees
       print(round(first_year_employees/total_employees,2))
```

```
0.28
```

28% of employees are in their first year of employement in the small charter school network.

```
[350]: #Creating column 'Tiers' based on the length of employment
       employment_length['Tiers'] = employment_length['Employment_Length_Overall'].
        ↪apply(lambda x: 'Tier 1' if x < 2 else ('Tier 2' if 2 <= x < 5 else 'Tier␣
        ↪3'))
```

```
[339]: #Creating subplot
       fig, axs = plt.subplots(nrows = 1, ncols = 2, figsize = (10,10))

       #Visulaizing Tier Metrics
       (round(employment_length['Tiers'].value_counts(normalize = True),3) * 100).
        ↪sort_values(ascending = True).plot(kind = 'pie', title = 'Employment Length␣
        ↪Tier Percentages', autopct= '%1.2f%%', ax = axs[0])

       #Visualizing Campus Median Length of Employment
       bar_chart = round(employment_length.
        ↪groupby('Campus')['Employment_Length_Overall'].median(),2).
        ↪sort_values(ascending = False).plot(kind = 'barh', title = 'Median Length of␣
        ↪Employment - Campus', figsize = (8,5), xlabel = 'Years', ax = axs[1])

       #Adding labels to visualization of Campus Median Length of Employment
```
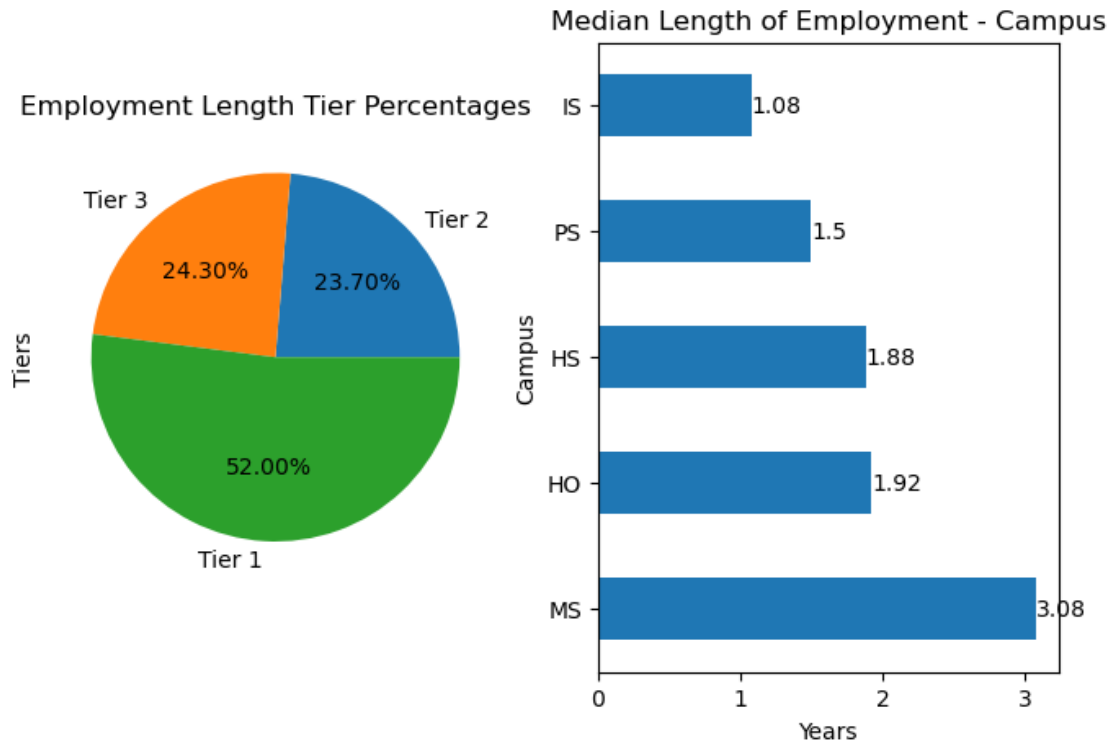
```
for i, v in enumerate(round(employment_length.
 ↪groupby('Campus')['Employment_Length_Overall'].median(),2).
 ↪sort_values(ascending = False)):
    bar_chart.text(v,i,str(v),va = 'center', color = 'black',fontsize = 10)
```


More than half of the employees in the small charter school network have been employed for less than or equal than 2 years. Additionally, the IS campus, also known as Upper ES, has the lowest typical length of employment, approximately 1 year 1 month, while the MS campus has the highest typical length of employment, 3 years and 1 month.

2. What is the difference in the projected and actual salaries for both departments and campuses?

[230]:
```
#Finding overall budget metrics
overall_budget = round(df[['Actual','Projected']].sum(),2)

#Finding difference between Projected and Actual Salaries
overall_budget['Difference'] = round(overall_budget['Projected'] -↩
 ↪overall_budget['Actual'],2)

#Creating variables for colors to assign to values
colors = ['Green','Orange','Blue']

#Creating variable to assign values in data labels
```
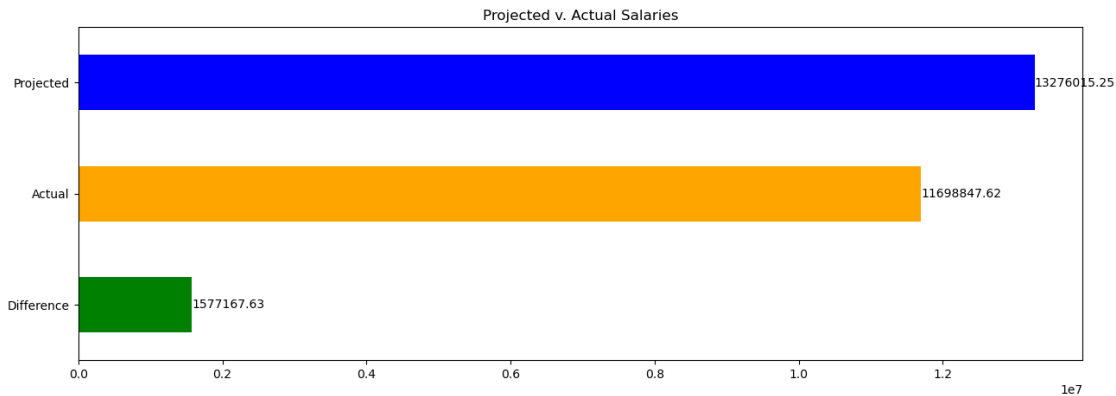
```
salaries_barh = overall_budget.sort_values().plot(kind = 'barh', title =␣
 ↪'Projected v. Actual Salaries', color = colors, figsize = (15,5))

#for loop to assign labels to bars
for i,v in enumerate(overall_budget.sort_values()):
    salaries_barh.text(v,i,str(v), va = 'center', color = 'black', fontsize =␣
 ↪10)
```

Projected v. Actual Salaries

Projected	13276015.25
Actual	11698847.62
Difference	1577167.63

Overall, there is a surplus of nearly 1.6 million USD in the budget.

```
[257]: #Creating df with departments and their projected and actual salaries
       department_overall_budget = df[['Department','Actual', 'Projected']]

       #Excluding the vlaues 'HS' & 'K-8' from the df 'department_overall_budget'
       department_overall_budget =␣
        ↪department_overall_budget[department_overall_budget['Department'] != 'HS']
       department_overall_budget =␣
        ↪department_overall_budget[department_overall_budget['Department'] != 'K-8']

       #Creating column 'Difference' by subtracting the 'Projected' and 'Actual'␣
        ↪columns
       department_overall_budget['Difference'] =␣
        ↪department_overall_budget['Projected'] - department_overall_budget['Actual']
```

```
[258]: #Calculating the Difference in the Department Projected and Actual Salaries
       department_overall_budget.
        ↪groupby('Department')[['Projected','Actual','Difference']].sum().
        ↪sort_values(by = 'Difference')
```

```
[258]:                          Projected      Actual  Difference
       Department
       Improvement of Instr. Svcs  160164.00  411948.00  -251784.00
       Business Office             332081.00  368123.70   -36042.70
```

```
Custodial                    738655.75    767770.00    -29114.25
Other Prof. Staff -CST       494495.00    516904.83    -22409.83
IT                           271234.00    274454.07     -3220.07
SST                          165165.00    166942.50     -1777.50
Transportation                74000.00     74000.00         0.00
Other - Instructional        195900.00    195615.25       284.75
Clerical Business             56250.00     54360.00      1890.00
Salaries (Support) - Nurse   261615.00    257006.25      4608.75
School Admin                1827599.00   1765513.03     62085.97
School Admin - Clerical     1844926.00   1712951.95    131974.05
```

The Department of Instructional Services has the largest defiict in the network, more than a quarter of a million dollars. The Department of School Administration - Clerical has the largest surplus in the network.

[285]:
```python
#Renaming columns
df.rename(columns = {'Actual_Budget_Salary':'Actual'},inplace = True)
df.rename(columns = {'Finance_Budget_Salary':'Projected'},inplace = True)
```

[291]:
```python
#Creating difference column
df['Difference'] = df['Projected'] - df['Actual']

#Calculating campus budgets and sorting by difference in ascending order
df.groupby('Campus')[['Projected', 'Actual','Difference']].sum().sort_values(by␣
 ↪= 'Difference')
```

[291]:
```
           Projected        Actual    Difference
Campus
HO        4255235.75    4272318.73     -17082.98
HS        2572115.00    2427286.50     144828.50
PS        2214785.50    1933109.24     281676.26
MS        2110753.00    1701112.90     409640.10
IS        2123126.00    1365020.25     758105.75
```

The HO Campus has the largest defiict in the network while the IS Campus, Upper ES, has a surplus exceeding three-quarters of a million dollars.

3. How many positions were unbudgeted and what is the amount of unbudgeted salaries?

[304]:
```python
#Creating a df with unbudgeted hires
unbudgeted = df[df['Budgeted']=='Unbudgeted']
```

[307]:
```python
unbudgeted.shape[0]
```

[307]: 11

[309]:
```python
round(unbudgeted['Actual'].sum(),2)
```
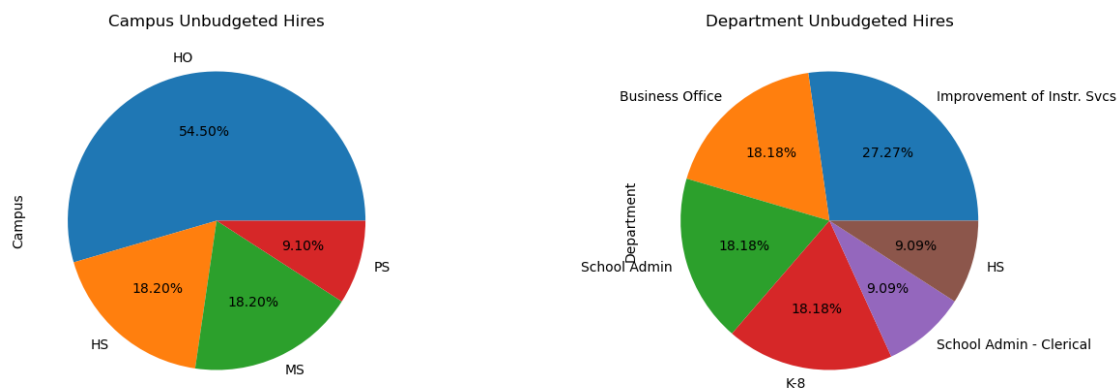
[309]: 878822.56

```
[333]: round(unbudgeted['Actual'].describe(),2)
```

```
[333]: count        11.00
       mean      79892.96
       std       28204.42
       min       19798.86
       25%       65100.00
       50%       74500.00
       75%       98525.00
       max      122500.00
       Name: Actual, dtype: float64
```

```
[326]: fig, axs = plt.subplots(nrows = 1, ncols = 2, figsize = (15,5))
       round(unbudgeted['Campus'].value_counts(normalize = True),3).plot(kind = 'pie',
        ↪title = 'Campus Unbudgeted Hires',ax = axs[0],autopct = '%1.2f%%')
       round(unbudgeted['Department'].value_counts(normalize = True),3).plot(kind =
        ↪'pie',title = 'Department Unbudgeted Hires', ax = axs[1],autopct = '%1.2f%%')
```

```
[326]: <Axes: title={'center': 'Department Unbudgeted Hires'}, ylabel='Department'>
```

There were 11 unbudgeted hires with salaries totaling 878,822.56 USD. More than half of the unbudgeted hires were at the HO Campus while approximately 27% of the unbudgeted hires were in the Department of Improvement of Instructional Services.

1.4 Conclusion

More than three quarters, 76%, of the staff in the small charter school network have been employed for 5 years or less. Approximately 28% of employees in the small charter school network are in their first year of employment at . More than half of the employees,52%, in the small charter school network have been employed for between two to five years, and the typical length of employment is approximately 2 years.

As it pertains to finances there is currently a surplus of approximately 1.6 million USD. The Department of Instructional Services has the largest deficit of approximately a quarter of a million

dollars while the Department of School Administration has the largest surplus of funds of approximately 132,000 USD. Home Office has the largest deficit of any campus of -17,082.98 USD. On the other hand Upper ES Campus has the largest surplus of all campuses of more than three quarters of a million dollars.