Looking at the Cost of IT for Hospitals via 990 Forms

By: Vidhi Patel

Data overview

- All the 990 forms where the **Taxpayer_name** contained the word "Hospital" for years 2011-2021 were downloaded.

- There were 26,404 entries however, some of them were empty or contained a different format:
  - Zero byte: 31 files
  - Version Not Supported: 1204 files
  - Data files: 25,168

- Hence for this analysis 25, 168 files were used.

Data Pre-processing

- New files were created that only included data for **Part IX:** Statement of Functional Expenses.
- Some files were then again dropped as they contained info for a different part IX and not the Part IX for functional expenses.
  - The total number of files became 23,802.
- Certain lines for Part IX, contained sub-categories as well.
  - For example, line 11: Fees for services had subcategories a - e.
  - I combined the values for such lines into 1 total cost.
- Some categories were also renamed to be more descriptive when graphing.

Final Dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 472179 entries, 0 to 472178
Data columns (total 7 columns):
 #   Column       Non-Null Count    Dtype
---  ------       --------------    -----
 0   object_id    472179 non-null   float64
 1   line_number  472179 non-null   object
 2   description  472179 non-null   object
 3   value        472179 non-null   float64
 4   perc_cost    469864 non-null   float64
 5   name         472179 non-null   object
 6   year         472179 non-null   int64
dtypes: float64(3), int64(1), object(3)
memory usage: 25.2+ MB
```

```
data['object_id'].nunique()
```
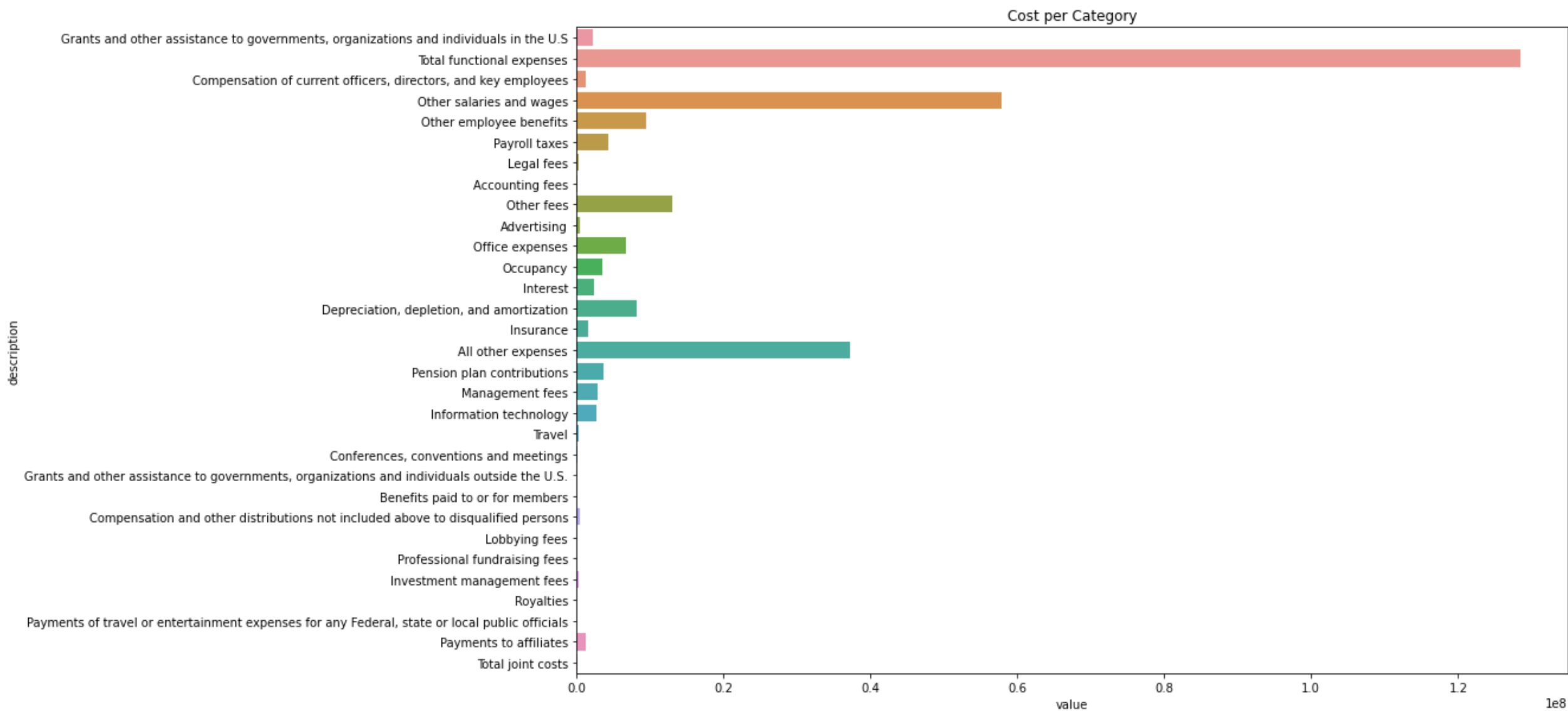
23802

```
data['name'].nunique()
```

5689

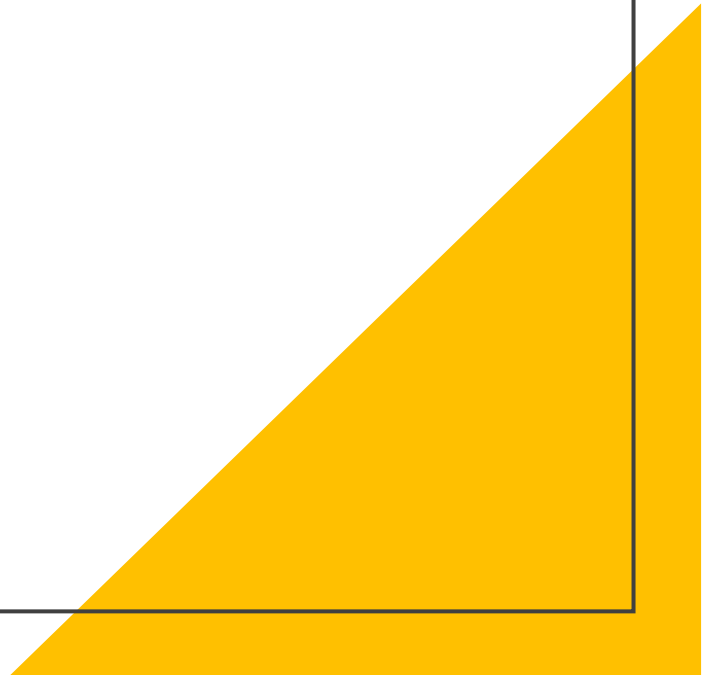The final dataset contains 472,179 entries with 23,802 unique records:

- Object_id: The unique id associated with the file
- Line_number: the corresponding line number in the 990 form
- Description: the description of the cost.
- Value: the cost in USD
- Perc_cost: the percentage of the cost of each description out of the total functional cost. This is calculated per object_id.
- Name: the name of the hospital
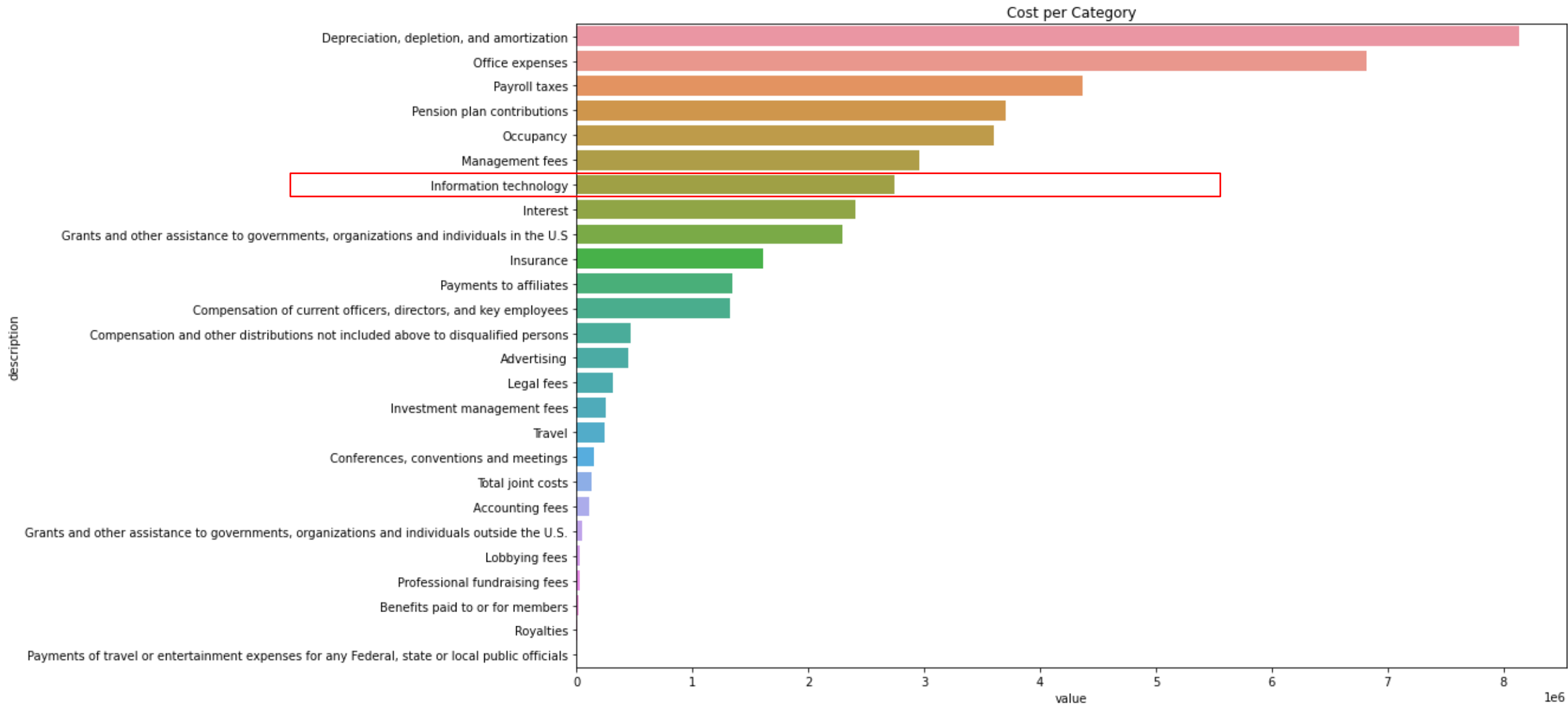- Year: The year of the collected data

The dataset contained 5,689 unique hospitals.
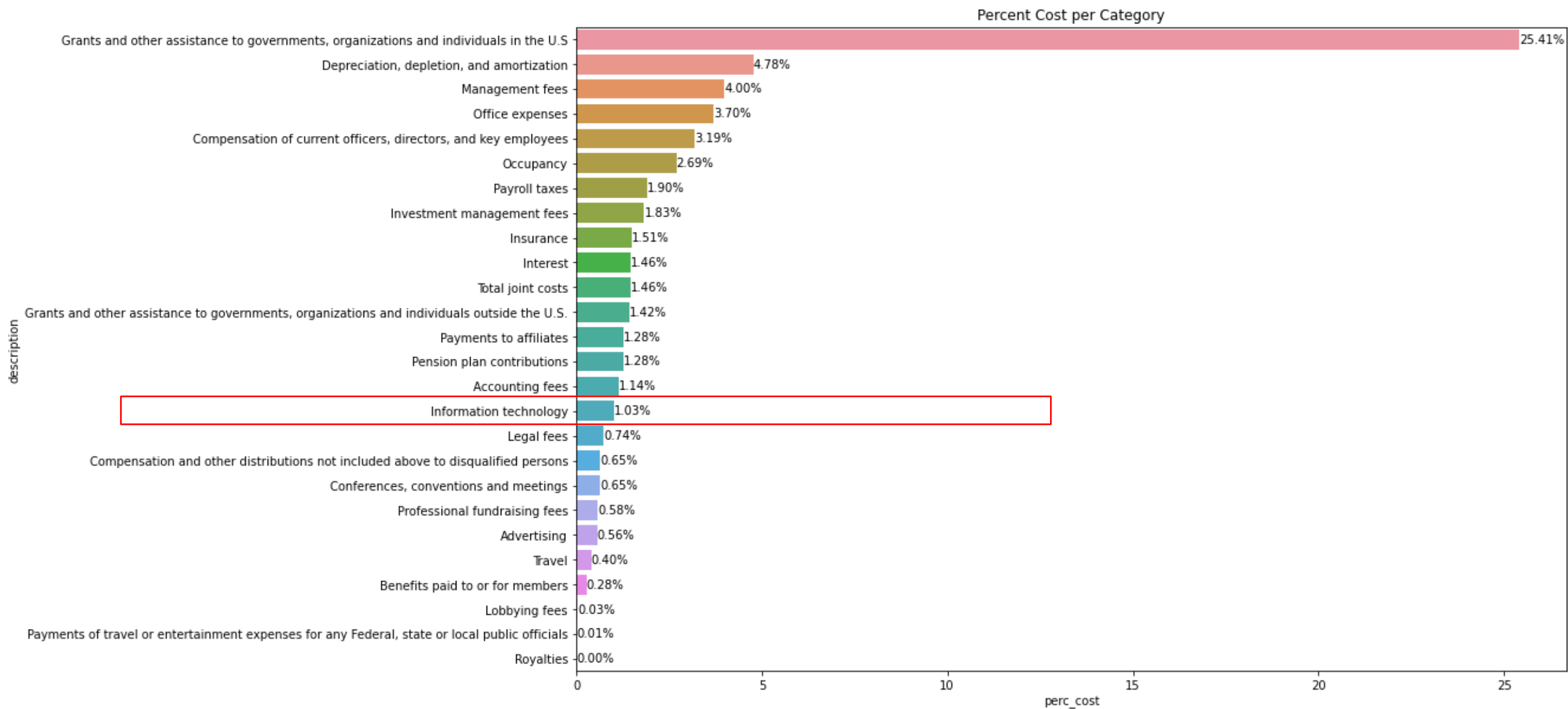
Cost per Category

Cost per Category

- As seen above, it is hard to see the cost for each category due to the "Total Functional Expenses" and other major contributions.
- The following descriptions were dropped as they are irrelevant to our goal:
  - Total functional expenses
  - Other salaries and wages
  - Other employee benefits
  - All other expenses
  - Other fees
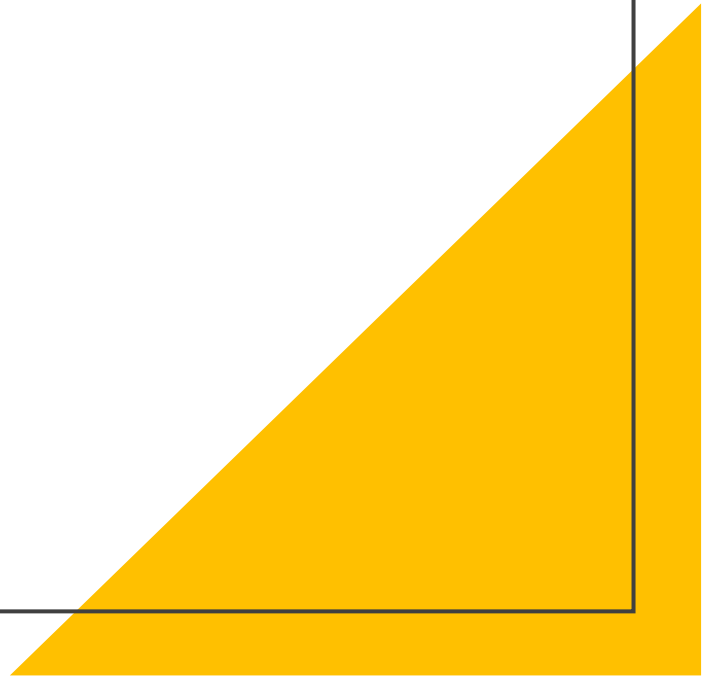- This gives us the plot below:

Cost per Category

This shows that IT is in the top 10 when looking at the cost in USD, however what percentage of the total cost is spent on IT?
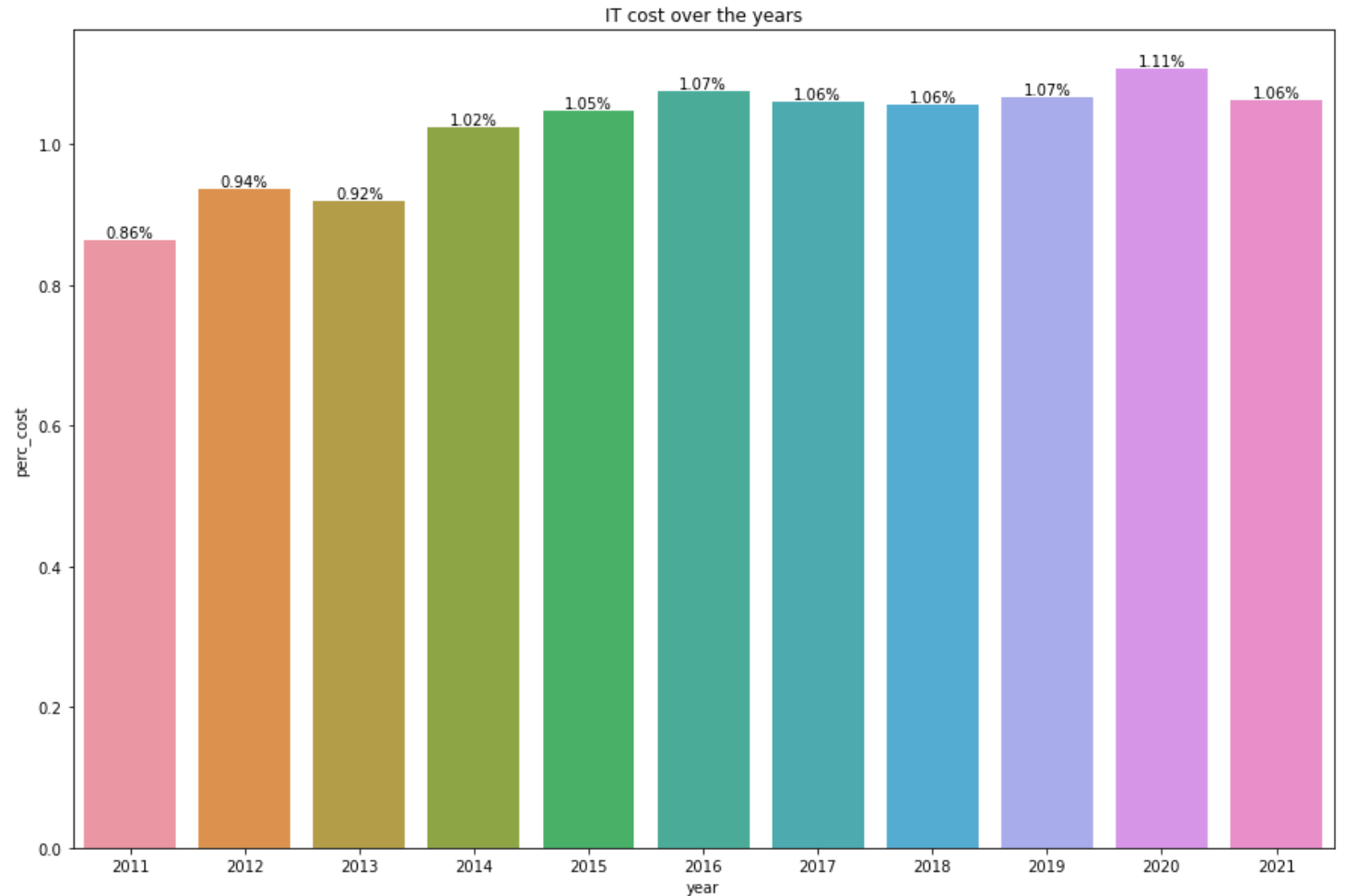
On average IT is only 1.03% of the total cost
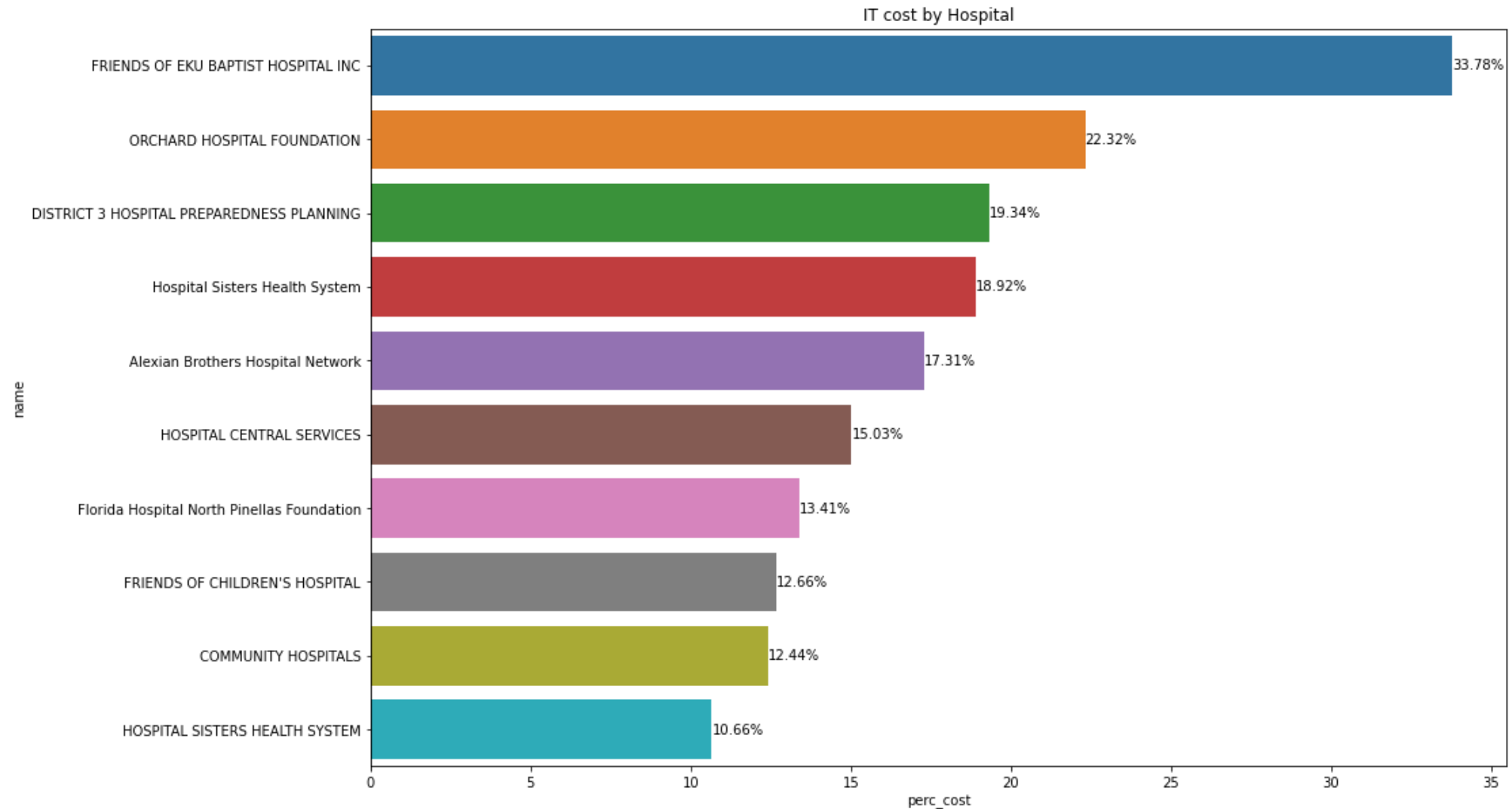
Looking at
IT Cost Only

IT Cost by Year
(percentage of total)

- As seen on the right, The IT cost has not significantly changed over the decade.

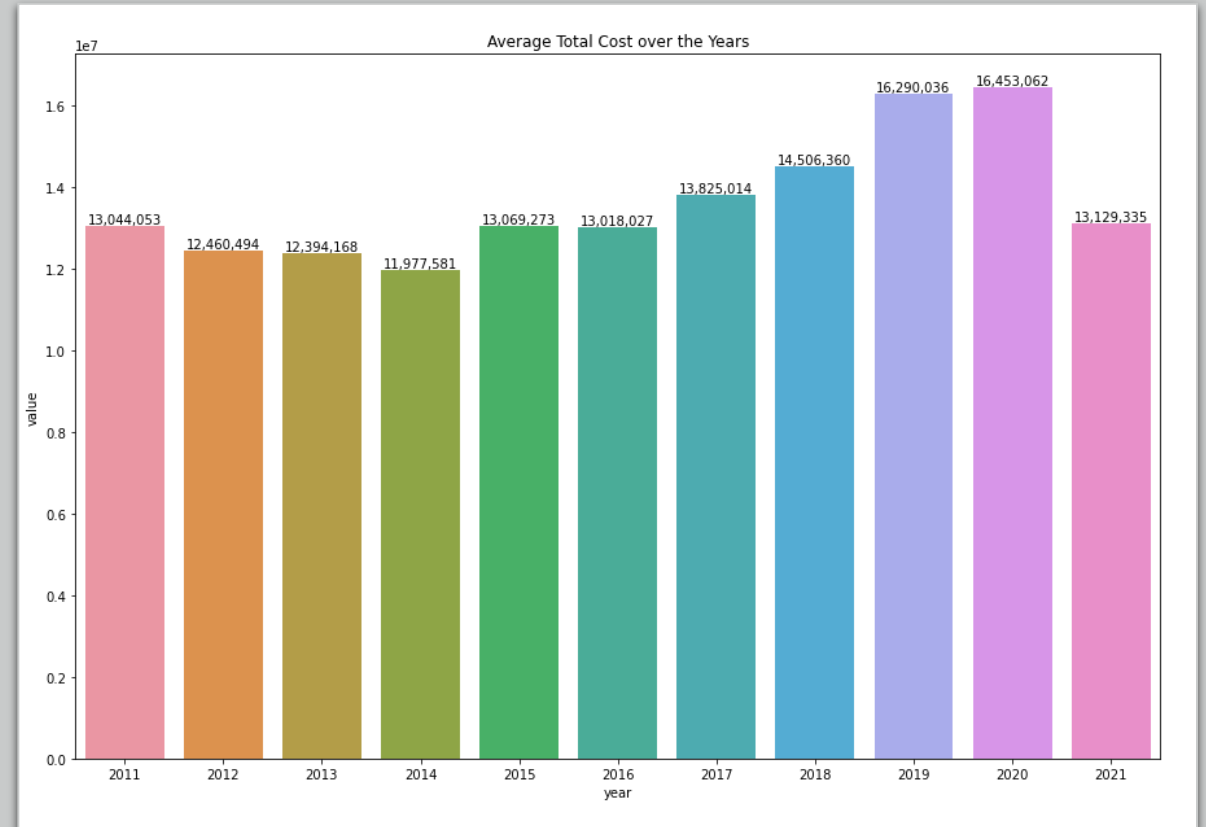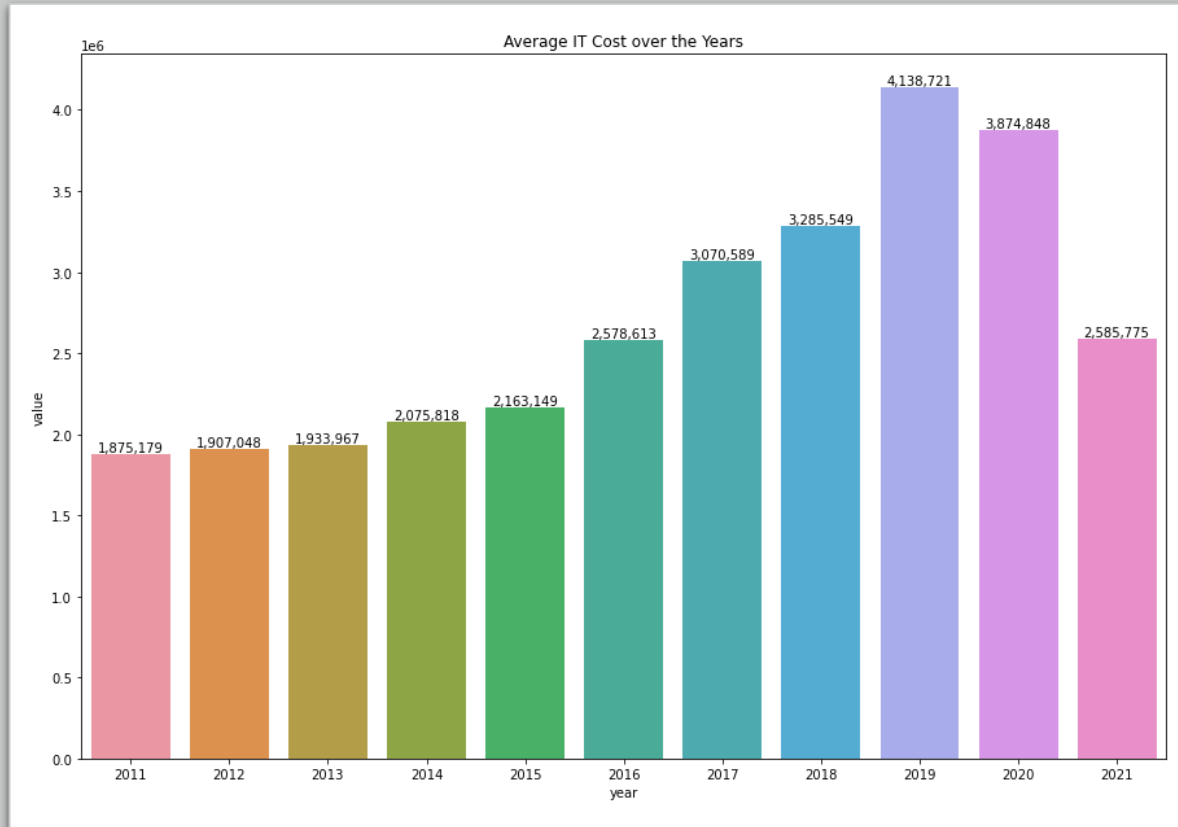- However, IT cost has increased by 0.2% from 2011 to 2021.

IT cost over the years

Top 10 IT cost by Hospital

Looking at the cost in USD

- The average cost of IT is increasing exponentially, however the total cost isn't increase that much over the years.

End.