problem-sheet-01
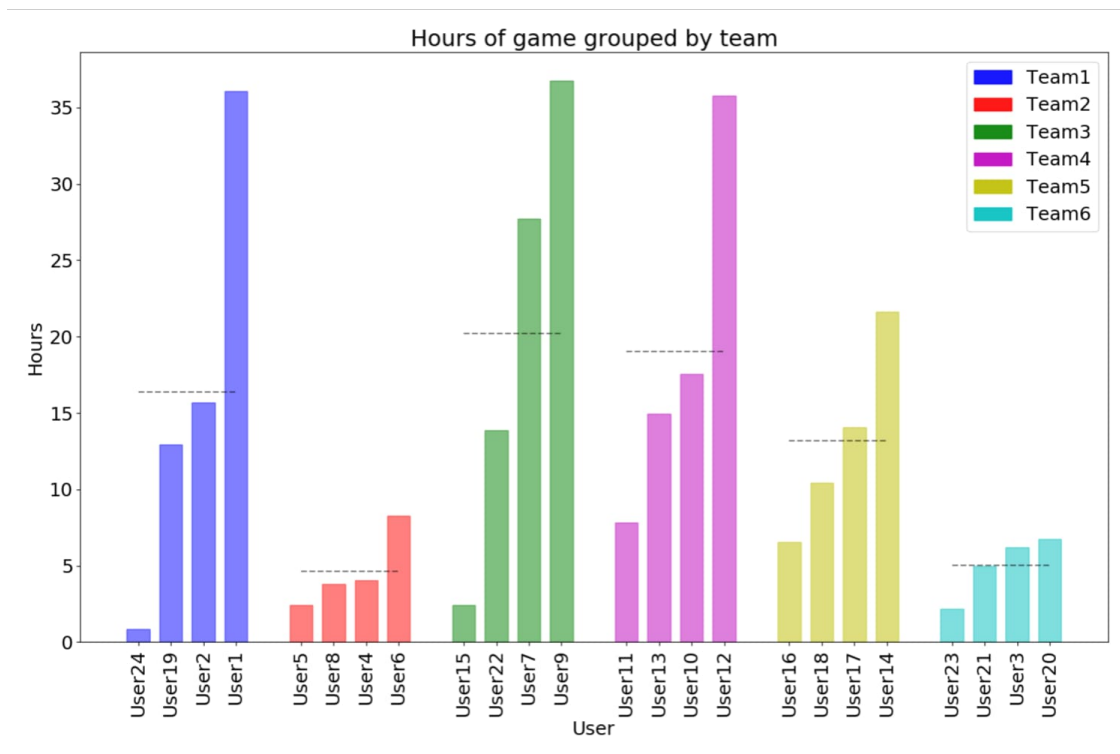
May 3, 2022

1 Problem sheet 1
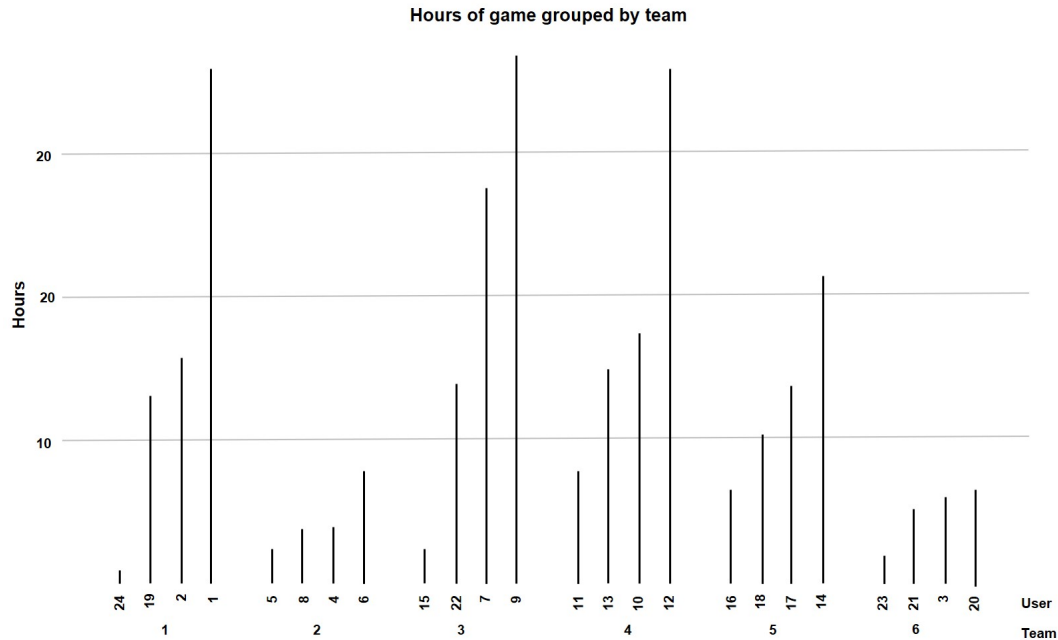
Hanne Hastedt (26014013), Jonas Lührs (23617618), Matteo Meier (21222337)

1.1 Exercise 1.1: Tufte and data-ink.

Original plot: (Source https://dl.acm.org/doi/10.1145/3139923.3139929)

Tufted plot:

For the tufted plot, we added grid lines for better readability. Eventually, they can be excluded from the plot as well.

Data-ink: labels, bars, legend (redundant), dotted line, axis, title

Non-data-ink: color, frame, width of bars

```
[1]:  # necessary imports
      import pandas as pd
      import plotly.express as px
      import matplotlib
```

1.2 Exercise 1.2: Visualizing the geography and economy of Countrystan

Data preview

```
[2]:  # load data from csv and show sample data
      df = pd.read_csv('data.csv', delimiter=' ', skiprows=1,
          names=['long', 'tlat', 'tagrGDP', 'ttier'])
      df['ttier'] = df['ttier'].apply(lambda x: str(x))
      df.sort_values(by='ttier', inplace=True)
      df.head(3)
```

```
[2]:          long        tlat   tagrGDP ttier
      49  7.497054  49.197906  1.000000     0
      25  5.572575  46.265576  0.332729     0
      26  3.264254  47.899908  0.237347     0
```
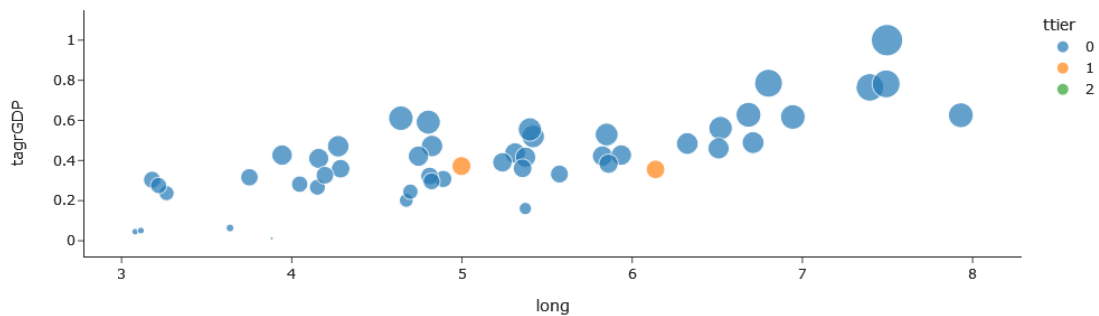
Plot No.1 (Longitude and Latitude at the axes):

2

```
[3]:   # scatter plot with long and tlat at the axes
       fig = px.scatter(df, x='tlat', y='long',
                        size='tagrGDP', color='ttier', template='simple_white')
       fig.show()
```
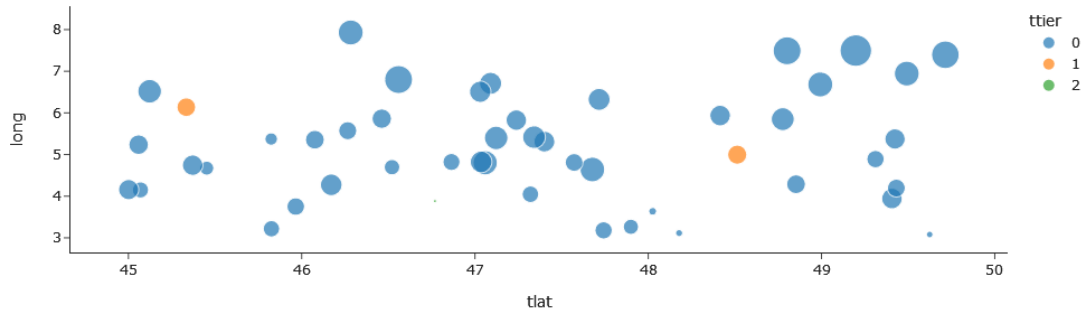

Plot No.2 (Longitude and GDP at the axes):

```
[4]:   # scatter plot with long and tagrGDP at the axes
       fig = px.scatter(df, x='long', y='tagrGDP',
                        size='tagrGDP', color='ttier', template='simple_white')
       fig.show()
```


The larger the longitude, the higher the tagrGDP. The capital with ttier=2 has a very small tagrGDP, one of the province capitals (ttier=1) has a tagrGDP=0, so it does not appear in the plots. The other two province capitals follow roughly the same pattern as the cities with ttier=0, but the province capital with smaller longitude has a higher tagrGDP than the other one.