Visualization Tools & Libraries

Pere-Pau Vázquez – Dept. Computer Science – UPC

Outline

- Required features

- Out-of-the-Box tools

- Libraries

Required features

- Different aspects:
  - Amount of programming
  - Static vs dynamic
  - Supported representations
  - Property/storage of the data!

Required features

- Amount of programming
  - [Almost] Nothing:
    - Out-of-the-Box tools, e.g. Tableau (bought by Salesforce, June 2019), Qlik, PowerBI, Plecto, Cumul.io, Looker (bought by Google, June 2019), Zoho Analytics
  - Medium:
    - Libraries over programming languages, e.g. JavaScript + D3 (with examples), Python + Matplotlib/Seaborn/Bokeh/Plotly/geoplotlib...
    - Custom highly flexible software: VEGA, Altair (over VEGA), Lyra,
  - High:
    - VEGA, Altair, JavaScript + D3, Python + X, C++ + Qt, R+ggplot2

Required features

- Importance of Visualization and Business Intelligence market
  - Logi Analytics buys Zoomdata (June 2019)
  - Salesforce acquired Tableau (June 2019) for 15.7B $
  - Google bought Looker (June 2019) for 2.6B $
  - Prezi buys Infogram (2017) for 70M
  - Palantir buys Silk (2016)
  - Mineset Inc. acquired by ESI Group (Feb 2016)
  - Stroër acquires Statista (Dec 2015) for 57.2M€
  - Microsoft bought Revolution Analytics (Jan 2015) and Datazen (Apr 2015)

Required features

- Static:
  - Cannot interact with the data
- Dynamic:
  - Allows for interactive selection, brushing, dynamic data changes…
  - Fundamental feature for data exploration

Required features

- Static:
  - Out-of-the-Box tools (PowerBI, Tableau, Excel, Charticulator…)
  - Python + Seaborn/cartopy/geopandas/matplotlib, R+ggplot2
- Dynamic:
  - Finished tools (limited)
  - JavaScript + D3
  - C++ + Qt
  - Python + Bokeh/Plotly
  - VEGA/Altair/Lyra…

Where is my data?

- Some solutions require to upload the data
  - Need to have permission
    - Medical data may be dangerous
  - Data may be sent outside your country
    - Check for example FaceApp application!
- SaaS may seize your data/visualizations
  - Visualization may be also out of your reach
    - If you cancel your subscription
    - If anything happens
      - E.g. the company files a bankruptcy, is acquired…

Discussion

- Supported representations:
  - Bar charts, scatterplots, pie charts… everywhere
  - Specific visualizations:
    - Maps
    - Trees
    - Networks
    - Text (word clouds, topics…)

Discussion

- Hybrid tools:
  - R + wrapper + plotly/highcharts/leaflet
- Programming languages/tools:
  - R+ggplot2: Good for statistical analysis, less optimal for nice visualizations, more complicated for interactive visualizations
  - Vega: grammar for graphics, somewhat similar to ggplot2
    - Vega-lite simplifies the grammar by taking several values by default
  - Python + matplotlib: similar to the previous ones, difficult for interactive visualizations

Discussion

- Programming languages/tools:
  - R+ggplot2: Statistical analysis
  - Python: Good for large data processing
    - Easy simple visualizations, more difficult to get interactive visualizations
    - Though some libraries exist
  - JavaScript: Easy in most devices (web)
    - Easy to tweak any visualization
    - Easy to implement interaction

Visualization Tools & Libraries

Pere-Pau Vázquez – Dept. Computer Science – UPC