HKUST
VISLAB

COMP 4462
Data Visualization Tutorial

Leo Yu Ho, Lo
Wenchao Li

Friday 17 April, 2019

https://bit.ly/vis-t07

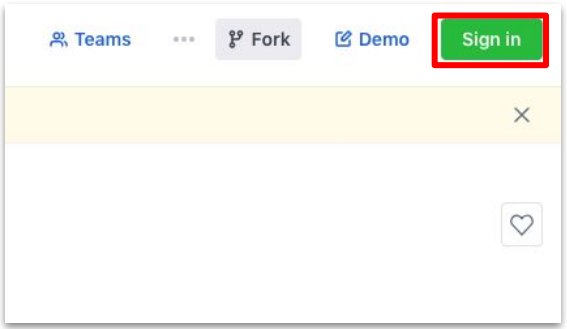
Vega-Lite and Data Processing Libraries

- Vega-Lite
  - The library behind Altair
  - Specification based visualization tool
    - We write down the visualization we want (in JSON format), the library plots it
    - In visualization language, marks and channels, interaction idioms, etc.
  - Builds on top of D3.js
    - And D3.js is build on top of HTML5 SVG (a web standard implemented in every browser)
  - See the Vega-Lite examples to know more
- Built-in functions in Javascript
  - Javascript borrows a lot of features from functional programming paradigm
    - Passing in a function as argument into another function
    - Makes our code much more succinct and easy to understand
- Lodash
  - An utility library for Javascript, a lot of common tasks and patterns are well written for use
- Moment.js
  - A powerful library for datetime manipulation

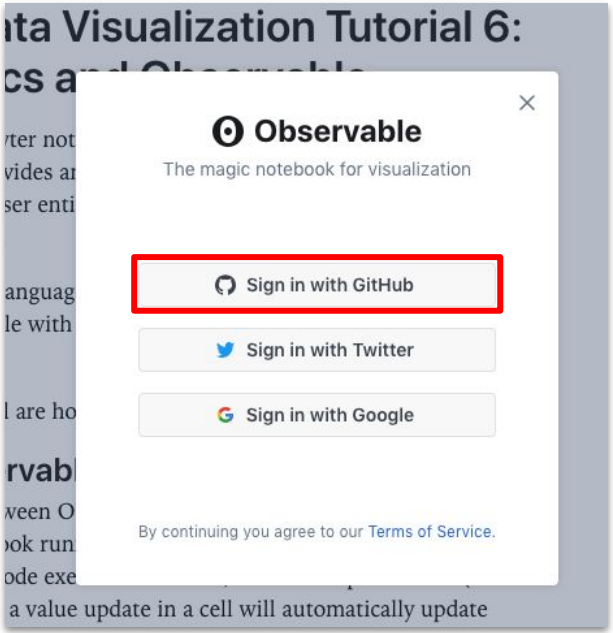
Sign in Observable

1. Go to the [notebook of this tutorial]

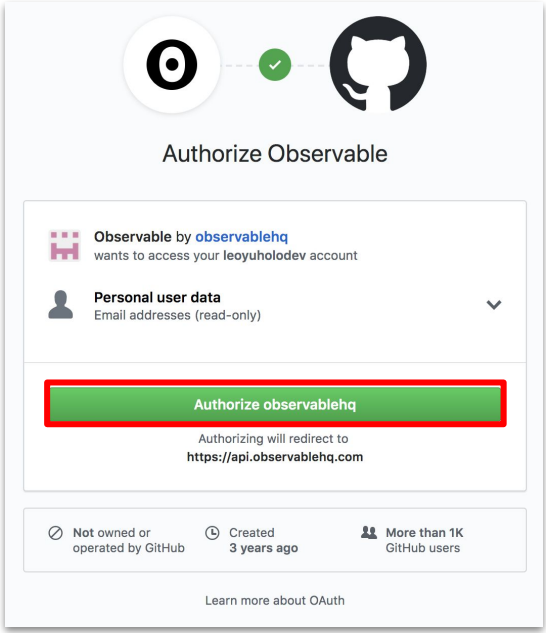
2. Click Sign in



3. Sign in with GitHub (recommended)



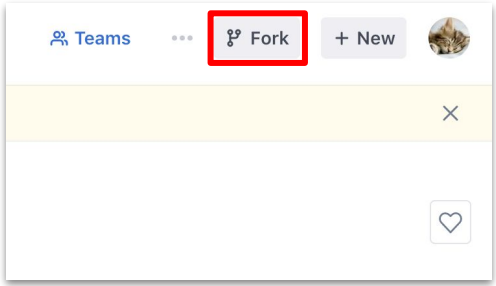
4. Authorize observablehq



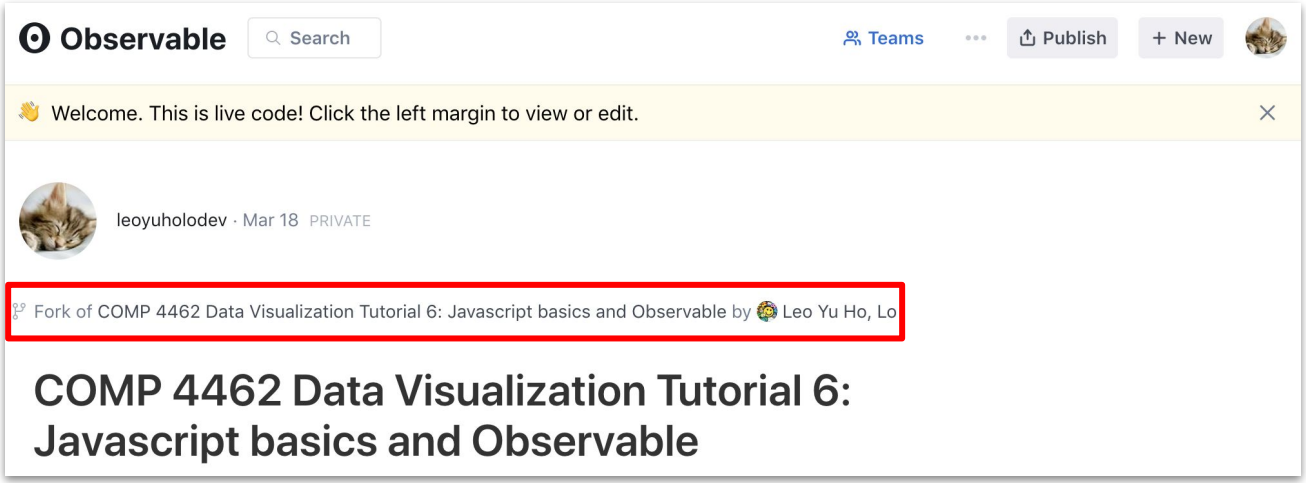
Fork Observable notebook

1. Go to the notebook of this tutorial

2. Click Fork



3. Check if you're working on your copy of the notebook (otherwise, your work will not be saved)



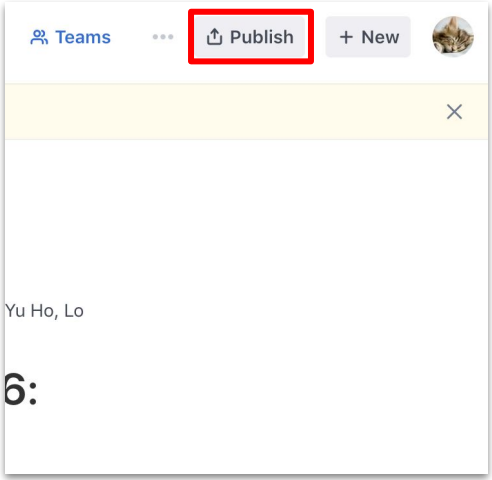
Data processing with Javascript

- See the Observable notebook of this tutorial
- Javascript built-in functions
  - map/reduce/filter
  - trim/split/indexOf/substring/replace
- Lodash
  - map over objects
  - groupBy / minBy / maxBy / meanBy
  - zip
- Moment.js
  - parse / format / datetime arithmetic
- Vega-Lite
  - Heatmap
  - Scale
  - Built-in aggregation
  - Datetime

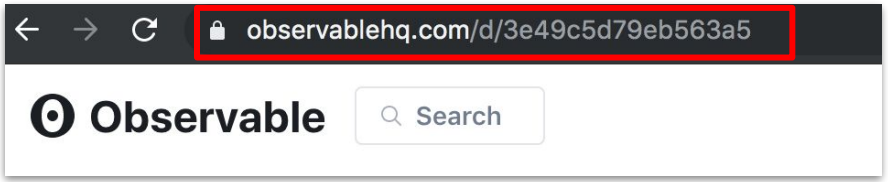
Publish your Observable notebook

1. In your working copy of the notebook

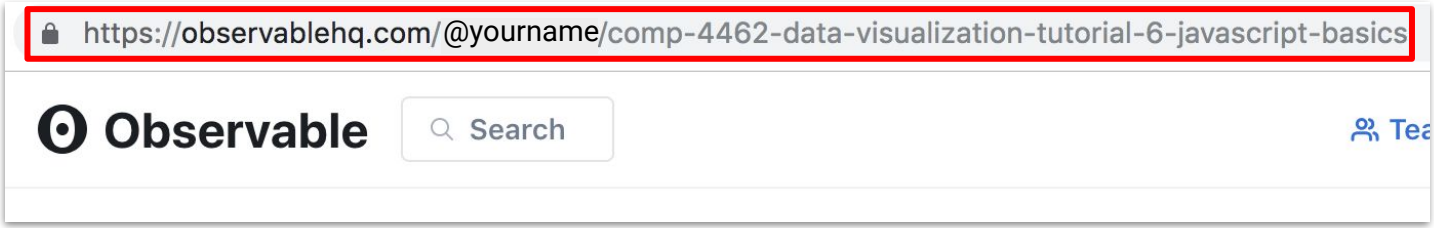
2. Click Publish



3. The URL changes from a hash...



4. To a meaningful URL, this URL is publicly accessible



Lab exercise

- Tasks
  - Sign in Observable
  - Open this Observable notebook and fork it (otherwise, your work will not be saved)
  - Read through “Data Processing with Javascript” and fill in the “TODO” cells
  - Prepare the Hong Kong temperature data from 1997 to 2017 for plotting
  - Plot heatmaps of the maximum/minimum temperature of each month
  - Use Vega-Lite built-in aggregation to plot the same heatmaps
    - Caution! This may hang your browser
  - Publish your notebook when finished
  - Copy the URL of your Observable notebook and submit to Canvas
    - The URL should be something like:
      - https://observablehq.com/@yourname/comp-4462-data-visualization-tutorial-7-vega-lite-and-data-p
- Optional
  - Star our GitHub repository ⭐⭐⭐ and like our Observable notebook ❤️❤️❤️ Thank you! ❤️
  - See more Vega-Lite examples and know more about what you can do with lodash and moment.js

More on Vega-Lite and data processing libraries

- More on Vega-Lite
  - View composition / layering / horizontal/vertical concatenate / interactions / zoom / filter / highlight / customize axis/ticks / maps (plotting geographic data)
- Notable functions of Lodash
  - sortBy / partition / transform / shuffle / sample / meanBy / sumBy / countBy / flatten / flattenDeep / mapKeys / mapValues / invoke / default / assign / merge / uniq / union / difference / repeat / deburr / split / words / chain
- More on Moment.js
  - Parsing and formatting / comparing / durations / handling timezone
- Other libraries:
  - apache-arrow: A future standard for in-memory data processing
  - JS libraries try to provide functionalities as Pandas to Python:
    - Data-Forge, Zebras, DataFrame-js

Next tutorial

Visualization with D3.js

- We will use Observable again
- And learn about D3.js!