Session 2

Data Visualization

Agenda

- A word about Jupyter
- Grading
- Presentation of Dash
- Playing with DASH

A word about Jupyter

- **I expect you are already familiar with Jupyter/Anaconda**
- **Usable in Final project**
- **Main issue for projects is libraries**
  - You may use any libs
  - I may not have them on my machine
  - Take the needed precautions
    - « plug&play » is expected

Jupyter online

Heard about Google Colaboratory ?

It allows to run your notebook in the cloud.

This will be accepted as medium for end of module project.

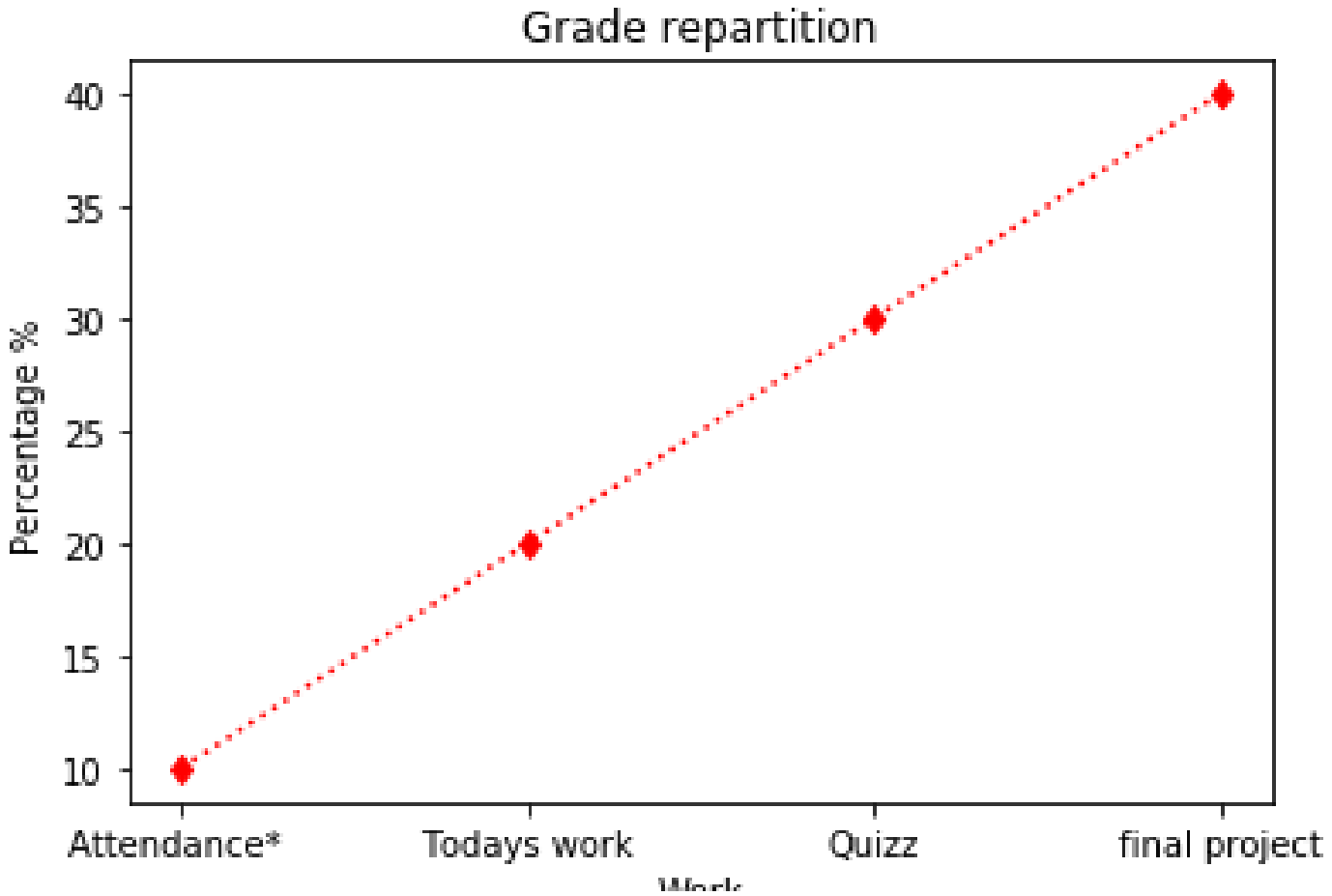
https://colab.research.google.com/notebooks/intro.ipynb

Grading

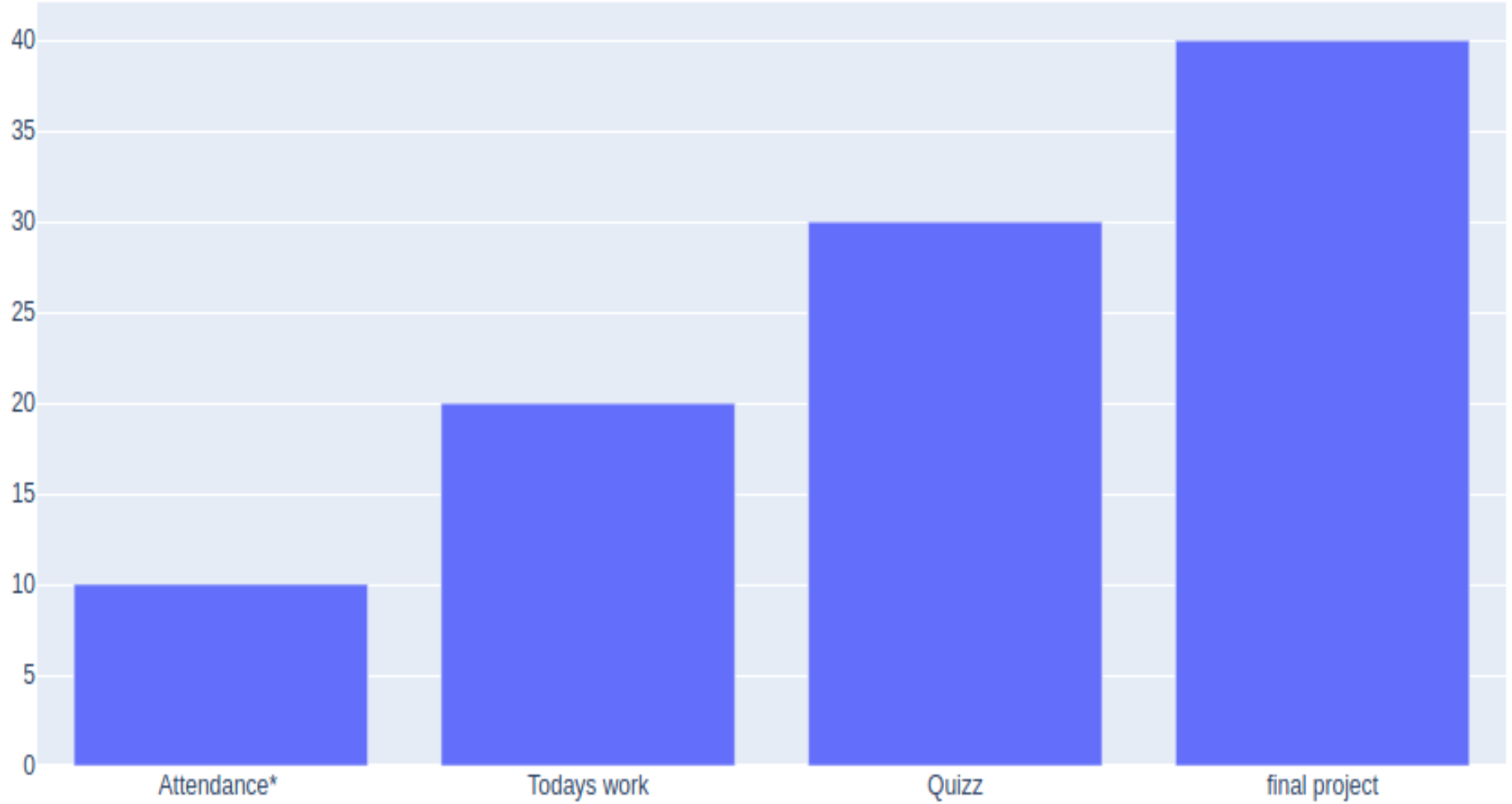
Grading for this module is divided like this

- **‘No issue’ attendance** **10 %**
- **Today’s work** **20 %**
- **Quizz** **30 %**
- **Final project** **40 %**

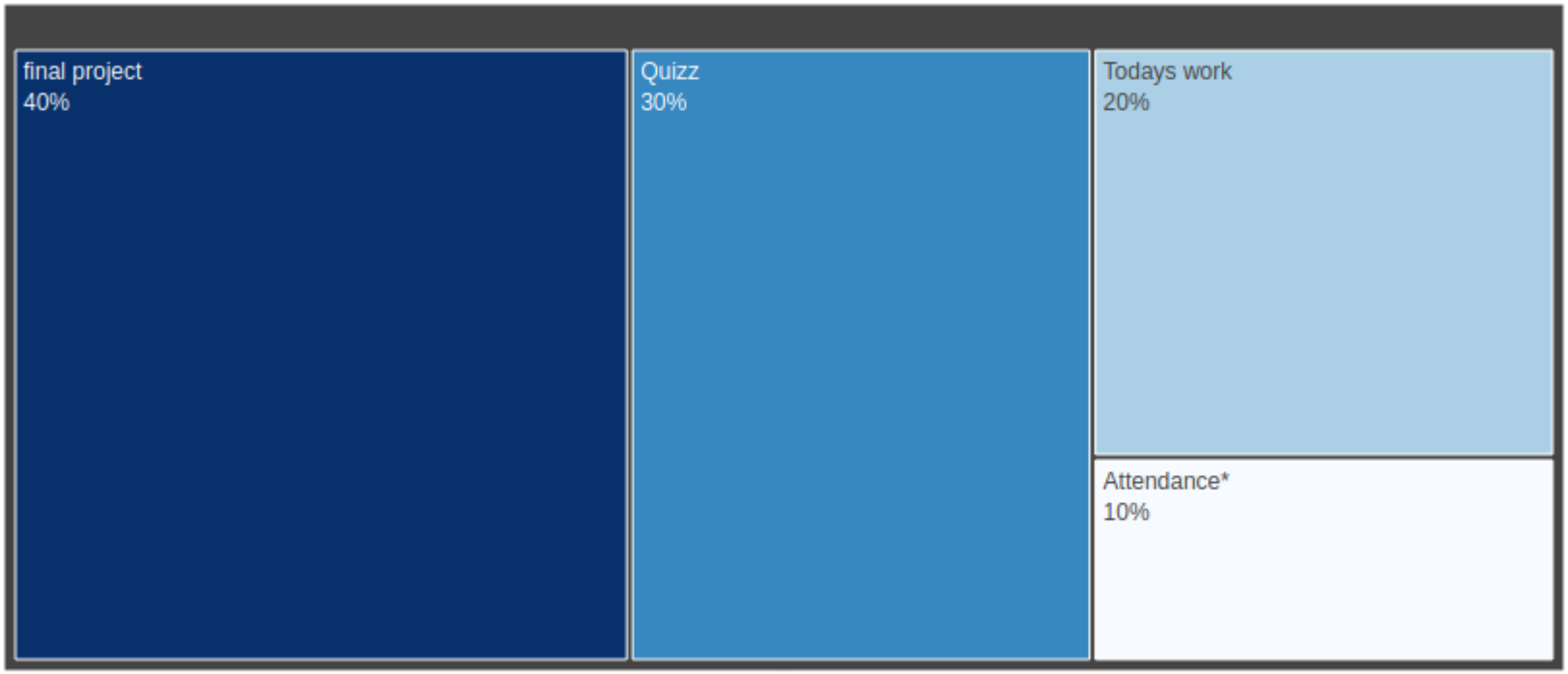
Same data, several dataviz



Same data, several dataviz



Same data, several dataviz



Same data, several dataviz

Code is in Grading.ipynb

Available in shared folder

Plenty of fish ...

There is a lot of Python libraries available

Matplotlib

Plot.ly <----------

Not only Python, same for Javascript

D3.js is amongst the most famous

Presentation of Dash

Today, we will see DASH

Dash is made by the team behind the plot.ly library

It allows to create (rather) quickly web apps that can show different kinds of charts (35 at this time)

URL is https://plotly.com/dash/

Comparable (in goal) to Shiny for R

Shiny is not authorized for EOM project

Gallery

Lets have a look at the example projects in the gallery at

https://dash-gallery.plotly.host/Portal/

Todays project is discovery of DASH and creating a dashboard.

Lets have fun

https://dash.plot.ly/getting-started

Why Dash ?

Besides what it is for and adding a tool to your toolbelt

It uses the plot.ly library, which you may already know as it can be used in Jupyter

There is also a DASH for R

Seems like a good choice for investing time in it !

Installation

I encourage you to use a Linux virtual machine

In which you can create a virtual environment.

sudo apt-get install python3-pip

sudo apt-get install python3-venv

virtualenv venv

source venv/bin/activate

Today’s project (20%)

Create a dashboard with DASH

Theme is your choice. I’ll validate or not.

No duplicates allowed.
First come, first served.

Produce insightful charts, of different natures and complexities.

Delivery .

How does it work ? The steps

Install

Create a .py file

Run this file

Example files were provided

Whats in a Dash file ?

- **Starts with imports**
  - Pandas, …
- **Part 1 Load data , arrange data**
  - Load CSV file(s) into dataframess
    - Clean data
- **Part 2 Create charts**
  - Know what you want to draw
    - Find the adhoc chart type with data-to-viz flowchart
    - Find the chart in the plotly documentation https://plotly.com/python/
    - Code your charts according to what we saw in session1
- **Part 3 Create a layout**
  - Start with a DIV
    - A DIV has children (consider as tags within the tag)
    - Populate it with an array of components
    - DIVs can contain DIVs
- **Part 4 Callbacks**
  - To discover later
- **Ends with a line to start the server (Flask)**

Structure - example

```
# -*- coding: utf-8 -*-
import dash
import dash_core_components as dcc
import dash_html_components as html
import pandas as pd
import plotly.express as px
external_stylesheets = ['https://codepen.io/chriddyp/pen/bWLwgP.css']

app = dash.Dash(__name__, external_stylesheets=external_stylesheets)

#  PART 1 -   load data (+arrange data)
df = pd.read_csv('https://raw.githubusercontent.com/plotly/datasets/master/2014_apple_stock.csv')
#  PART 2 - Produces charts
fig = px.line( ....

#  PART 3 - Create layout

app.layout = html.Div(children=[
    html.H1(children='Hello Dash'),
    html.Div(children='''
        Dash: A web application framework for Python.
    '''),
dcc.Graph(
        id='aapl-graph',
        figure=fig
    )

])
#  PART 4 - Callbacks (later)
# end of part 4
if __name__ == '__main__':
    app.run_server(debug=True)
```

This example is simplified.

it will not run

20 % Dash app : Expectations

Are expected today :

- **A running app**
- **Several charts of several kinds**

Are not expected today :

- **Live charts with selectable values (see Dash callbacks)**
- **Tabs (Single page showing different charts)**

Expectations

- **You choose a subject+find a dataset**
  - The quicker the better
- **I validate it or not**
- **A running app is expected at 13h**
  - Containing several charts of several kinds (minimum is 3)
  - Delivered in a zip, by email
    - Zip to be named after
      - Your group
      - Your name
    - Put those informations in the page too