GFT

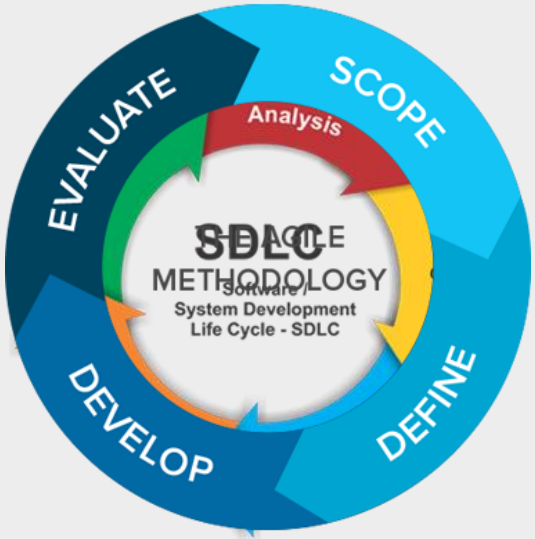
Notebooks

Pedro Nieto, 03/04/20

Agenda

1. **What is a Notebook**
2. **Different Notebooks**
   - Jupyter
   - Zeppelin
3. **Best Practices**
4. **Zeppelin**
   - Elements
   - Interpreters
   - Usage

What is a Notebook? First current state..



How long does a cycle take?

What if I change my mind?

What if I want to Access real data?

Agile?

What is a Notebook?

- But what happens when I want to execute code in Production Directly?

Scenario 1:

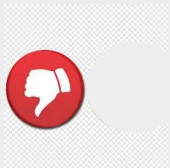
“ I need to run a query in Production”

PL/SQL IDE 

Scenario 2:

“ I need to run Python Code in Production”

Python IDE 



What is a Notebook?

Under some circunstancies running code in production directly is needed:

- Production Support
- Reporting
- Analysis
- Data Science

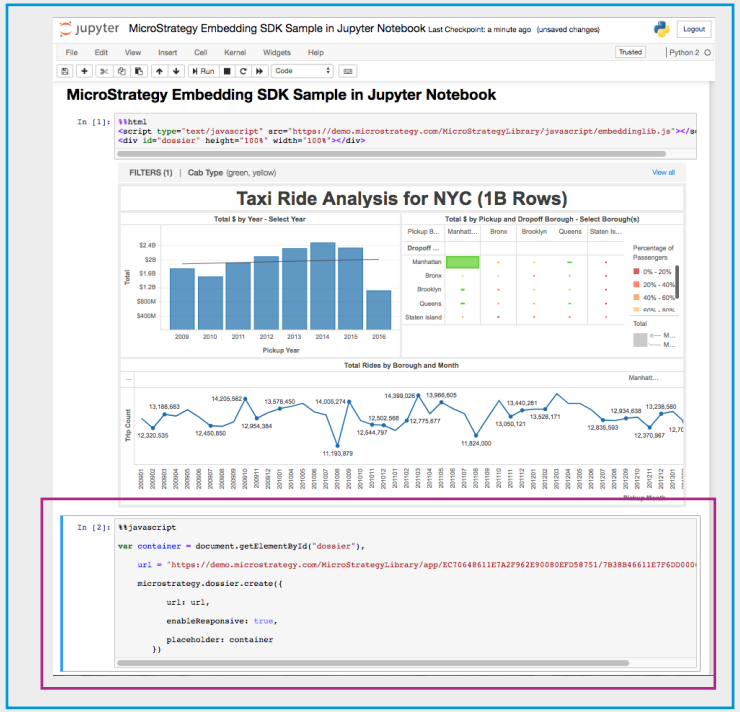
How do we solve this?

What is a Notebook?

An **open-source** *web application* that allows you to create and share documents that **contain live code**, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

The idea is providing the users the tools needed for rapid prototyping, data analysis… without the need of waiting for a development lifecycle.

What is a notebook? Parts…



Facts:

- The server that runs the Notebook is called **Kernel**
- Variables are **by default shared** between cells
- Variables are removed when kernel is restarted
- Cells can be executed **independently** and without order
- Notebooks can be **multi-language**, different code is called interpreter

Different Notebooks

1. Open collab
2. Log in with your Google Account
3. Create a new notebook
4. Create a paragraph with:
    1. print(“Hello World”)
5. Insert Snippets and execute them

1. Open Jupyter
2. Create a new notebook
3. Create a paragraph with:
   print(“Hello World”)
4. Copy and snippet from Google and execute: Is it working?
   Visualization: Interactive Scatter Plot in Altair
5. Create a new paragraph with the following
   !pip install vega_datasets
6. Execute paragraph with code again: Is it working?
7. Create a paragraph with:
   A = 5
8. Create a paragraph at the top with: Does this work?
   print(A)
9. Restart kernel and execute only 9 step paragraph. Working?

Best Practices

- #1 One notebook, one focus.
- #2 State is explicit.
- #3 Notebook is not an IDE
- #4 Tidy up your code
- #5 Label your diagrams

Apache Zeppelin

Let’s talk about Zeppelin!

Interface Parts

- General
  - Notebooks
  - Credential
  - Repo
  - Helium
- Notebook Management:
  - New
  - Import
  - Jobs

- Notebook
  - Options
    - Export
    - Versions
    - Visualization Mode
    - Interpreter Binding
  - Editor
    - Interpreter
    - Output
    - Width
    - Height
    - Font Size
    - Graphs

GFT

Shaping the future of digital business

Pedro Nieto

Data Specialist

Data / AI & IoT Practice Member

GFT Valencia