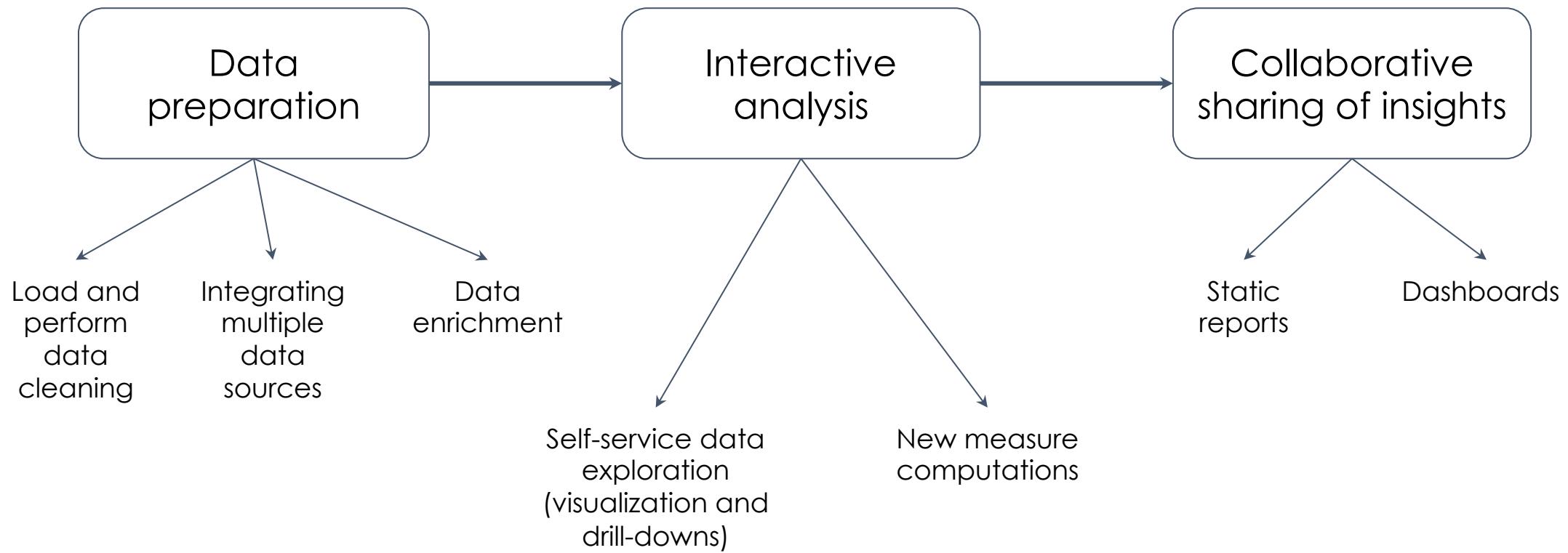


# Create BI analytics platform collaboratively on Google Cloud

Presented by Yeo Huifang

# What is a BI analytics platform

It is a self-contained architecture that enables non-technical users to execute full-spectrum analytic workflows from:



# Use Case Demo



- Advantage of technical resources
  - CPU / memory limitation
  - network connectivity
  - storage constraints
- Single source of truth
  - centralized source of data
  - all queries are read-only mode
  - ability to add new data source without restarting cube
- Collaborative sharing of insights
  - new data or measures will be available to all
  - immediate validation of analysis
  - non-technical users can interactively create their dashboards

# Platform deployment



## Tools Used

---



+

**atoti.**<sup>TM</sup>

+



**Model prototyping and experimentation**

**In-memory data cube and BI web application**

**Online sharing**

# Modes of deployment

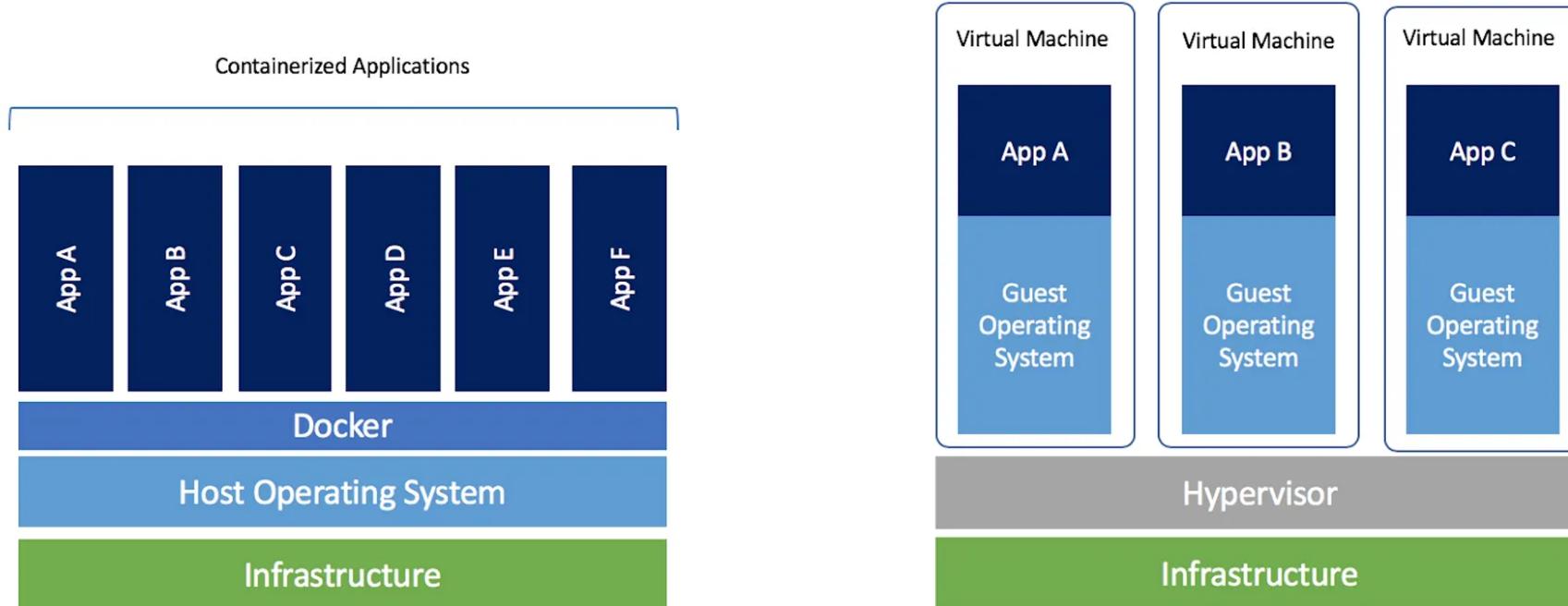


Image from [Docker Blog](#)

- Google Cloud Run
- AWS Fargate
- Azure container instance

- Google Compute Engine
- AWS EC2
- Azure virtual machine

- Get started for free with Google with \$300 free credit  
<https://cloud.google.com/free>
- Create a project
- Create a virtual machine
- Configure firewall (port 8888 for JupyterLab, 9999 for BI Web application)
- Install python (through pipenv or download conda)
- Install JupyterLab with atoti
- Configure JupyterLab
- Launch JupyterLab

<https://gist.github.com/HuifangYeo/1f9dd417246e66f6c4868ffa02fc961>

```
debian_python_setup.sh
Raw

1  # 1 - update system packages
2
3  sudo apt update -y
4
5  # 2 - install pyenv's dependencies
6
7  sudo apt install -y make build-essential libssl-dev zlib1g-dev libbz2-dev libreadline-dev libsqlite3-dev wget curl llvm libncurses5-dev lib
8
9  # 3 - install pyenv to install python on persistent home directory
10
11 curl https://pyenv.run | bash
12
13 # Configure the environment by adding pyenv to path
14 echo 'export PATH="$HOME/.pyenv/bin:$PATH"' >> ~/.bashrc
15 echo 'eval "$(pyenv init -)"' >> ~/.bashrc
16 echo 'eval "$(pyenv virtualenv-init -)"' >> ~/.bashrc
17
18 # updating bashrc
19 source ~/.bashrc
20
21 # install python 3.9.2 and make default
22 pyenv install 3.9.2
23 pyenv global 3.9.2
24
25 # execute
26 pip install --upgrade setuptools pip
27 pip install atoti[jupyterlab]
28
29 # install graphviz for more features with atoti
30 sudo apt install graphviz -y
```

- Get started for free with Google with \$300 free credit  
<https://cloud.google.com/free>
- Create a project
- Create a virtual machine
- Configure firewall (port 8888 for JupyterLab, 9999 for BI Web application)
- Install python (through pipenv or download conda)
- Install JupyterLab with atoti
- Configure JupyterLab
- Launch JupyterLab



- Create a free account with AWS  
<https://aws.amazon.com/free>
- Checkout [atoti.io](#) website > videos

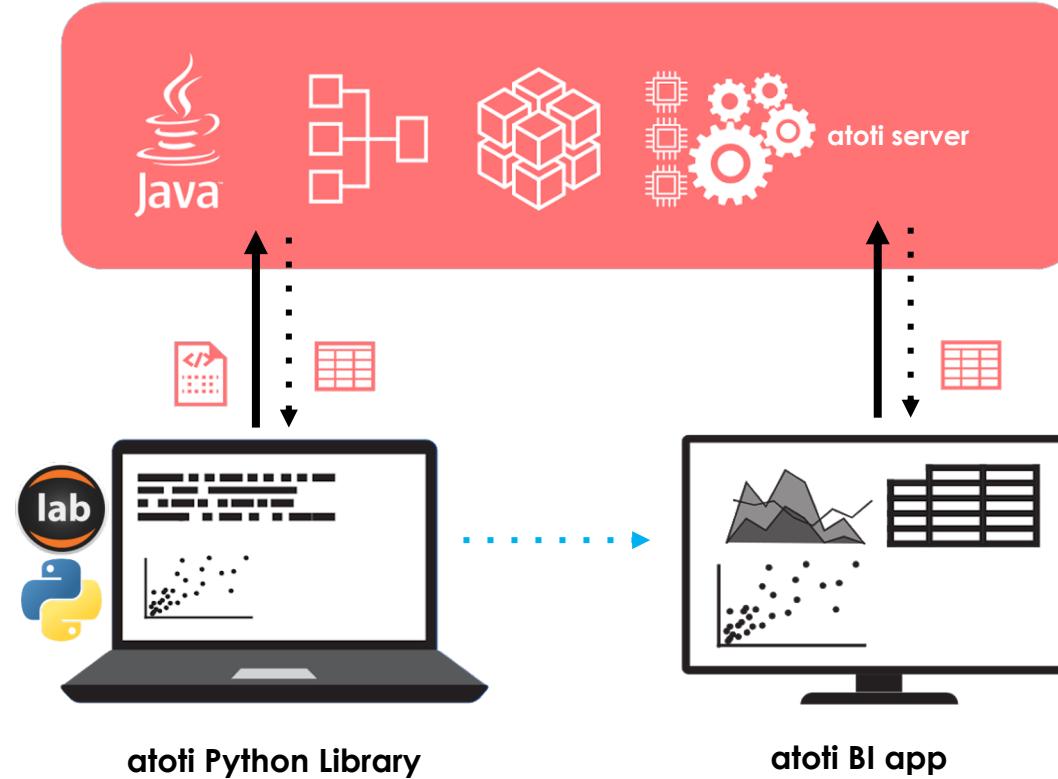
Docker: [Deploying atoti Docker image on AWS with Amazon ECS](#)

VM: [Deploying atoti on AWS with Amazon EC2](#)

What is atoti



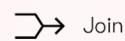
## Defining atoti



## ETL

### Optimized data loading

#### Schema Definition



Join



Type inference

#### Connectors



CSV



Pandas



Parquet



NumPy



SQL



Spark

#### Streaming & Incremental Updates



Kafka



Append



File watcher

#### Cloud Storage



AWS



Google Cloud Storage



Azure Blob Storage

## Aggregation

### Fast in-memory calculations

#### Arithmetic

Count, sum, product, distinct  
count, long, short, square sum...

#### Analysis

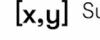


Min, max, arg min, arg max...

#### Statistics

Variance, standard deviation,  
mean, median, quantile...

#### Array



Sum, top, bottom, sort...

#### Mathematical

Exponential, power, logarithmic,  
root, round, floor, ceiling, cosinus,  
sinus, tangent...

#### Time

Datediff, rolling sum over a period,  
year-to-date, month-to-date...

## Analysis

### Efficient & resilient analysis

#### What-If Analysis

Scenario creation: reclassification,  
source simulation, measure scaling

Side-by-side comparison



Impact analysis

#### Multidimensional Analysis



Root cause analysis

Operations: slice and dice,  
drill-down, drill-up, roll-upFunctions": filter, where, parent  
value, shift, at, siblings...

Hierarchical data structure

#### Data Science

Translate model result into  
business KPIsModel result analysis: perfor-  
mance, bias...

## Data Visualizations

### Interactive BI

#### Widgets



Pivot table



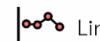
Combo chart



KPIs



Pie chart



Line chart



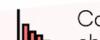
Scatter plot



Bar chart



Gauge



Column chart

#### BI Features

Story-telling  
analysis

Dashboard

Drag & drop  
interactive  
analysis

Collaboration

#### Visualization Tools



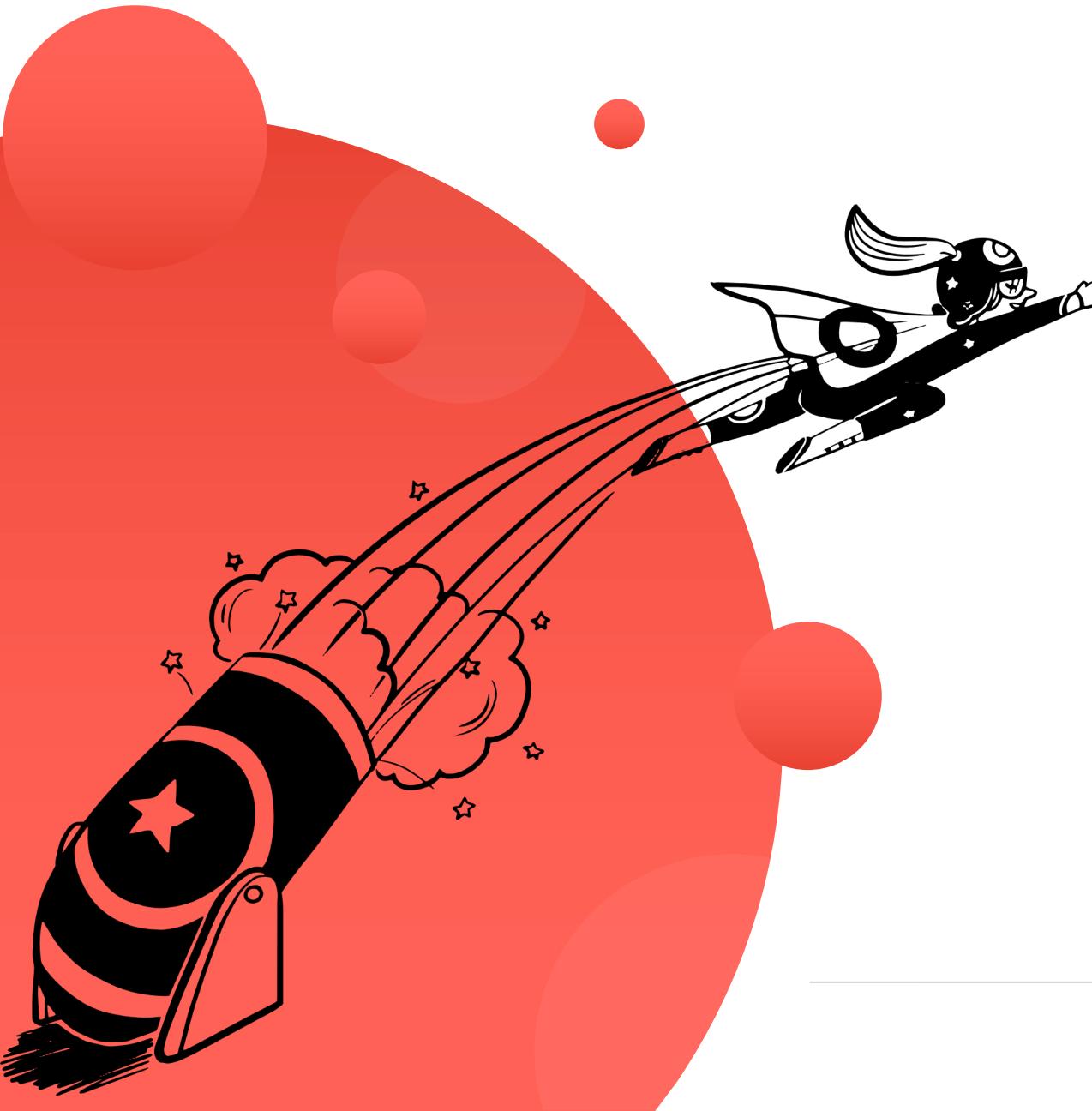
BI web application



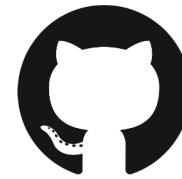
JupyterLab extension



Excel (XMLA support)



# Any questions?



- Name one expected functionality from a BI analytics platform
- Name one advantage of having an analytics platform on the cloud