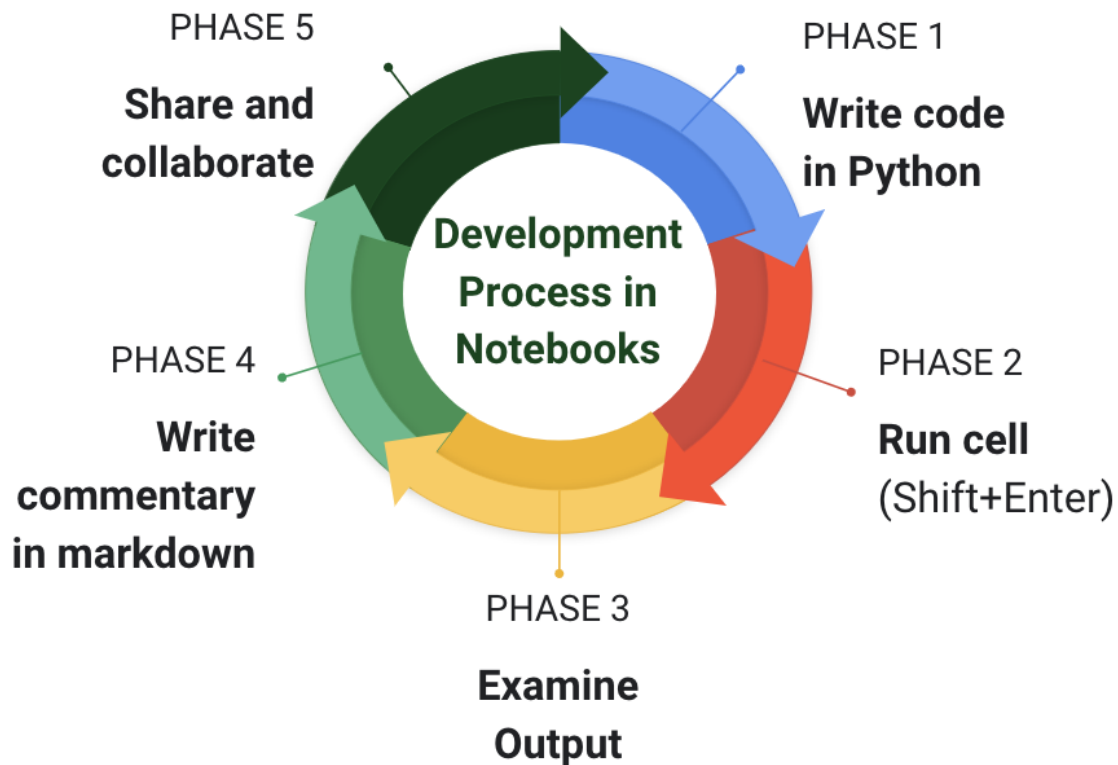AI Platform Notebooks
Besides BigQuery, the other piece of technology you will use in the first lab is AI Platform Notebooks. AI Platform Notebooks is the next generation of hosted notebook on GCP, and has replaced Cloud Datalab. This is a managed Jupyter notebook environment that you can use to run Python code. It handles authentication to GCP, so that you can easily access BigQuery.

PHASE 5
Share and collaborate

PHASE 1
Write code in Python

Development Process in Notebooks

PHASE 4
Write commentary in markdown

PHASE 2
Run cell
(Shift+Enter)

PHASE 3
Examine Output

AI Platform Notebooks are developed in an iterative collaborative process. You can write code in Python, hit the run button, and the output shows up right on the page itself. Along with the code, you can write commentary in Markdown format and share the notebook with your colleagues.

This is what the interface looks like. Notice how there are code sections interleaved with markup and output. This interleaving is what makes this style of computing so useful. Data analysis and machine learning are commonly carried out in notebooks like this. The code is in the blue section. You can execute the code by either clicking the Run button or by pressing Shift + Enter. The red section displays output from the command. Notice that the output here isn't just command line output; it's charts and tables as well. The yellow section contains markup, so you can explain why you're doing what you're doing in plain-text.
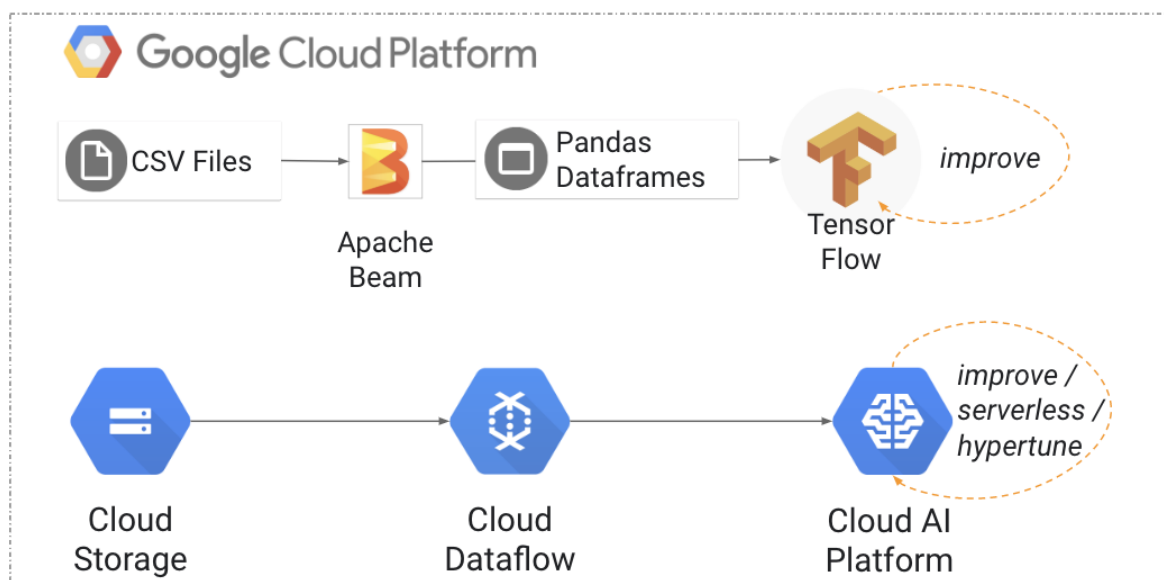
AI Platform Notebooks work with the same technologies that you're comfortable with, so you can start developing now, and then work on scale later. For example, we'll be doing an exercise where we read from a .csv file. You could then process in Pandas and Apache Beam before training a model in TensorFlow, and then improve the model through training.

Eventually though, when you are ready to scale, you can use Google Cloud Storage to hold your data initially, process it in Cloud Dataflow on an ephemeral cluster, and then run distributed training and hyper-parameter optimization in Cloud AI Platform.