Holly Figueroa

DSC650 Week1

PySpark example run results

```
!apt-get install openjdk-8-jdk-headless -qq > /dev/null
!wget -q http://archive.apache.org/dist/spark/spark-3.1.1/spark-3.1.1-bin-hadoop3.2.tgz
!tar xf spark-3.1.1-bin-hadoop3.2.tgz
!pip install -q findspark
```

```
import os
os.environ["JAVA_HOME"] = "/usr/lib/jvm/java-8-openjdk-amd64"
os.environ["SPARK_HOME"] = "/content/spark-3.1.1-bin-hadoop3.2"
```

```
import findspark
findspark.init()
from pyspark.sql import SparkSession
```

```
%run pi.py
```

```
    Pi is roughly 3.144680
```

Your URL should open in a new window. If it does not, make sure that pop ups are not blocked and reopen the link. ✕

Colab paid products - Cancel contracts here

Your URL should open in a new window. If it does not, make sure that pop ups are not blocked and reopen the link.