CODE BASE BY CHAPTER

2.2 Get the Data Server!

LINK: https://data36.com/digitalocean

2.3 Log in to your data server (for Mac/Linux users)

```
ssh root@[your_ip_address]
```

3.1 Install bash - and create a new user!

```
echo 'Hello, World!'
adduser [newusername]
adduser dataguy
usermod -aG sudo [newusername]
exit
ssh [newusername]@[your_ip_address]
```

3.2 Install Python 3 and Jupyter Notebook!

```
python3
print ('Hello, World!')
exit()

sudo apt-get update
sudo apt-get install mc
sudo apt-get -y install python3-pip
sudo apt-get -y install python3-dev
sudo -H pip3 install --upgrade pip
sudo -H pip3 install jupyter
jupyter notebook --generate-config

echo "c.NotebookApp.ip = '*'" >>
/home/[your_username]/.jupyter/jupyter_notebook_config.py

echo "c.NotebookApp.allow_remote_access = True" >>
/home/[your_username]/.jupyter/jupyter_notebook_config.py

jupyter notebook --browser any
```

LINK: [IP Address of your remote server from the email]:8888

3.3 Install postgreSQL!

```
sudo apt-get install postgresql postgresql-contrib
sudo -i -u postgres
psql
CREATE USER [your_user_name] WITH PASSWORD '[your_preferred_password]';

\q
exit

psql -U [your_user_name] -d postgres

CREATE TABLE test(column1 TEXT, column2 INT);
INSERT INTO test VALUES ('Hello', 111);
INSERT INTO test VALUES ('World', 222);
SELECT * FROM test;
\q
```

3.4 Install pgadmin4!

```
sudo -i -u root
echo "listen_addresses = '*'" >> /etc/postgresql/*/main/postgresql.conf
echo 'host all all 0.0.0.0/0 md5' >> /etc/postgresql/*/main/pg_hba.conf
sudo /etc/init.d/postgresql restart
Exit
SELECT * FROM test;
```

3.5 Install R and RStudio!

```
sudo apt-get install r-base-core
R
print ("Hello, World!");

quit()
sudo apt-get install gdebi-core
wget https://download2.rstudio.org/rstudio-server-1.0.136-amd64.deb
sudo gdebi rstudio-server-1.0.136-amd64.deb
sudo restart rstudio-server
```

LINK: [IP Address of your remote server from the email]:8787

4.2 Installing Python libraries: Pandas, NumPy, MatPlotLib and Scikit-Learn

```
sudo -H pip3 install numpy
sudo apt-get install python3-pandas
sudo -H pip3 install --upgrade beautifulsoup4
sudo -H pip3 install --upgrade html5lib
sudo -H pip3 install --upgrade scipy
sudo -H pip3 install scikit-learn
```

```
import numpy as np
import pandas as pd
import matplotlib
matplotlib.use('Agg')
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn.linear_model import LinearRegression
```

```
df = pd.DataFrame({'a':[1,2,3,4,5,6,7],
'b':[1,4,9,16,25,36,49]})
df.plot()
```

4.3 Installing csvkit

```
sudo -H pip3 install csvkit
```

+ Sources and further reads

Bash/Command Line:
Data Science At The Command Line: http://datascienceatthecommandline.com
R:
https://support.rstudio.com/hc/en-us/articles/200552306-Getting-Started
SQL:
https://www.cyberciti.biz/faq/howto-add-postgresql-user-account/
https://help.ubuntu.com/community/PostgreSQL
http://stackoverflow.com/questions/1287067/unable-to-connect-postgresql-to-remote-database-using-pgadmin
Python:
http://jupyter.readthedocs.io/en/latest/install.html
https://www.digitalocean.com/community/tutorials/how-to-set-up-a-jupyter-notebook-to-run-ipython-on-ubuntu-16-04