**Currently working with**

R

POWER BI - <https://www.ktlsolutions.com/bi/power-bi-r-marketing-campaign-analysis/>

SQLite

**Immediate to do list -> move this onto Asana or Trello**

Initial Database Work (free)

* GCP big query (yep lots of free interaction)
* AWS SQL (1 day instance)
* MongoDB (atlas is starting point – free stuff on AWS, Azure or GCP)

Want to just mess around with free s3

**By Type**

**Python**

Introduction to programming in Python on edx

Intro to Python for data science on datacamp

https://docs.python.org/2/tutorial/

http://scipy-lectures.org/index.html

**SQL**

https://www.datacamp.com/courses/intro-to-sql-for-data-science

<https://www.codecademy.com/learn/learn-sql>

SQL syntax similar but queer planner and parametrisation might change from  
 program to program (DB to DB)

Postsql – garbage collector

**AWS**

<https://www.aws.training/LearningLibrary?filters=digital%3A1>

Amazon aurora (SQL)

Amazon DynamoDB

**ML**

https://scikit-learn.org/stable/tutorial/basic/tutorial.html

https://ml.berkeley.edu/blog/2016/12/03/github/

https://github.com/eriklindernoren/ML-From-Scratch/blob/master/README.md#polynomial-regression

https://yerevann.com/a-guide-to-deep-learning/

http://neuralnetworksanddeeplearning.com/chap1.html

**JavaScript**

<https://www.w3schools.com/js/js_examples.asp>

**R**

https://www.edx.org/course/analyzing-big-data-with-microsoft-r

<https://www.edx.org/course/introduction-to-r-for-data-science>

https://www.datacamp.com/courses/free-introduction-to-r

**Hadoop**

**https://www.tutorialspoint.com/hadoop/index.htm**

**SQL**

scanning vs seeking

applying workarounds to queries

https://www.w3schools.com/sql/sql\_join.asp

**NO-SQL key value stored**

Key is very important.

JSON, BSON

MAP REDUCE

Unstructured -> map to new structure, reduce.

**More on Python**

Representational State Transfer (REST)service? Exposes data content

Rest is moving away from object orientation – HTTPs is not an object, state behaviour and data should be separated.

Read about rest python

https request – rest – abstract data store away

JSON customer

lambda calculus, functions

https://www.datacamp.com/courses/intermediate-python-for-data-science

**Postman**

Test http requests (lambda???)

**Elastic Search and Kibana**

Yes, Kibana is a free, open-source visualization tool. You can run Kibana on-premises, on Amazon EC2, or on Amazon Elasticsearch Service. With on-premises or Amazon EC2 deployments, you are responsible for provisioning the infrastructure, installing Kibana software, and managing the cluster. With Amazon Elasticsearch Service, Kibana is deployed automatically with your domain as a fully managed service, automatically taking care of all the heavy-lifting to manage the cluster.

***Large databases difficult to scale well into cloud***

**But…**

**Aws/azure/google cloud**

Infrastructure as service vs Platform as a service

Make VMs and install Cassandra onto it

Cloud is elastic.

Try to

Create a database in aws/azure

AWS EC2 – rest service, image. Pre0configure EC2 server.

Stateless AWS lambda – just a function … spins up when required.

**All Links**

introduction to programming in Python on edx

Intro to Python for data science on datacamp

quantum world on EDX which is mostly python but slightly more advanced

https://www.codecademy.com/

<https://www.edx.org/course/introduction-computer-science-mitx-6-00-1x-10>

<https://www.edx.org/course/introduction-computer-science-harvardx-cs50x>

<https://projects.iq.harvard.edu/stat110/about>

<https://www.coursera.org/learn/machine-learning>

<https://www.kaggle.com/>

<https://www.cs.ubc.ca/~murphyk/MLbook/>

<http://scikit-learn.org/stable/>

https://www.tensorflow.org/tutorials/

<https://github.com/eriklindernoren/ML-From-Scratch>

<http://interactivepython.org/runestone/static/pythonds/index.html>

<http://danluu.com/programming-books/>

<http://course.fast.ai/>

<http://yerevann.com/a-guide-to-deep-learning/>

<https://ml.berkeley.edu/blog/2016/11/06/tutorial-1/>

<https://github.com/jostmey/NakedTensor?bare>

<http://aiplaybook.a16z.com/>

<https://aws.amazon.com/blogs/aws/new-aws-training-and-certification-portal/>

<https://github.com/pauli-space/foundations_for_deep_learning>

<https://news.ycombinator.com/item?id=14486657>

<https://www.youtube.com/playlist?list=PLZ9qNFMHZ-A4rycgrgOYma6zxF4BZGGPW>

<http://cs109.github.io/2015/>

<https://distill.pub/>

3blue1brown deep learning

<http://neuralnetworksanddeeplearning.com/chap1.html>

**Recommended by others:**

[**https://www.udemy.com/**](https://www.udemy.com/)

https://docs.google.com/presentation/d/1n2RlMdmv1p25Xy5thJUhkKGvjtV-dkAIsUXP-AL4ffI/edit?usp=drivesdk

**MVP process:**

- mindset vs research

- how quick you can deliver some value to project now vs try to deliver the perfect value a year later,

*https://medium.com/idealo-tech-blog/what-is-minimum-viable-data-product-49269e338d85*

**General Resources:**

- some articles from

*https://towardsdatascience.com/*

**- podcasts:**

*https://lineardigressions.com/*

*https://www.dataengineeringpodcast.com/*

**SQL(you can get some experience using )**

bigquery in google

**ML**

Tensorflow - so many deep learning options -

pytorch

**AWS**

- Sagemaker - becoming all the rage with Data Science people,

https://docs.aws.amazon.com/sagemaker/latest/dg/gs-setup-working-env.html

- S3,

- EC2( just their barebones servers )

**GCP**

- colab - https://colab.research.google.com/notebooks/welcome.ipynb

- bigquery - they have some open datasets that you can get to grips with larger datasets and queries

**Visualisation –**

* python plotly   
  https://plot.ly <https://dash.plot.ly/>

SQL through AWS

<https://aws.amazon.com/getting-started/tutorials/create-mysql-db/?sc_channel=em&sc_campaign=global_F90D_DF_F7D_E2_tutorials_2017&sc_medium=em_57591&trk=em_57591&sc_content=adopt_f90d_f90d_ot&sc_geo=mult&sc_country=global&sc_outcome=adopt_f90d&mkt_tok=eyJpIjoiTnpVNVpUa3hPV1ZsTUdRNSIsInQiOiI4TmVjZXpZdXRtVGcxbHhcL2wya21tc20rcVVQUU5qNlhMa05cLzFtalkyMkFnakw0VUozeU9UcDZLMzF5Mk5rZGN4Z25vdHBlS2ZWaWhad2lXVlV0Y2xqd2pxZXhabXp2V2tcL3JjYk5YcjVhVDAyZktNc0VqWk9janpcLzFScjZLY3kifQ%3D%3D>

<https://dev.mysql.com/downloads/file/?id=484382>

Sever Through AWS