Jacob’s To-Do: 2213276

**Contents**

[Big Data: 1](#_Toc161045576)

[What do I have? 1](#_Toc161045577)

[What do I need? 1](#_Toc161045578)

[Machine Learning: 1](#_Toc161045579)

[What do I have? 1](#_Toc161045580)

[What do I need? 1](#_Toc161045581)

[Ruskin Module: 1](#_Toc161045582)

[What do I have? 1](#_Toc161045583)

[What do I need? 1](#_Toc161045584)

[Applied Machine Learning Project: 2](#_Toc161045585)

[What do I have? 2](#_Toc161045586)

[What do I need? 2](#_Toc161045587)

# Big Data:

## What do I have?

* The Datasets to use
* Some example commands
* Some useful reading links
* A MongoDB Atlas account
* A corresponding database cluster
* The virtual machine
* Ollie & Hugh as resources

## What do I need?

* To get the virtual machine working
* To run Compass in the VM
* To import the Data into the VM then into the Cluster
* To do stuff with the Data in the Cluster to meet Task 1 (80%) Criteria
* Prepare the presentation.
* Optionally do Task 2 (20%)

# Machine Learning:

## What do I have?

* The Brief & initial code provided on Canvas

## What do I need?

* To pick a ML model
* Write the jupyter notebook code

# Ruskin Module:

## What do I have?

* Nothing

## What do I need?

* To pick a Topic
* Decide on an Artefact
* Make the Artefact
* Write reflective writing on Artefact (~1000 words)

# Applied Machine Learning Project:

## What do I have?

* Gmail Test-Account
* Google Dev thingy related to api app permissions
* Basic code to access the Test-Account via the API
* Code Implementation of a NLTK tutorial, resulting in my own tokeniser code.

## What do I need?

* Code to process the tokenised output
* To carry on following current project plan
* Prep Logbooks for submission
* Continue Writing Report

## What am I doing?

### Week 10:

---> Bag of Words scikit-learn example implementation

---> ~~Email summation~~

---> ~~Email Categorization~~

---> Collate Paper Notes

--->