Goals:  
1) Set up Git Main and Branches

2) Share Ideas

3) Setting roles between members

7:32 pm Jerry first Idea

* The Great Pacific Garbage Patch, they accomplished the start of a clean up last week that removed a lot of the plastic, proving it can be cleaned up. Ive saved some articles ill share tomorrow. Just an idea but thought it might be interesting as its topical/current and probably has a ton of data connected

7:42 pm Created repo <https://github.com/alecngai/Data_Capstone_2021>, added members

# Project Ideas

## Jerry

* The Great Pacific Garbage Patch, they accomplished the start of a clean up last week that removed a lot of the plastic, proving it can be cleaned up. Ive saved some articles ill share tomorrow. Just an idea but thought it might be interesting as its topical/current and probably has a ton of data connected

## Alec

Track global mining hashrates and energy consumption,

Batch Machine Learning vs Live Machine Learning model

Node sql to temp store inputs,

Crypto Mining,

Machine learning aspect:

* Predicting hashrates?
* Correlation with energy costs and crypto mining density

Ref

<https://cbeci.org/mining_map/methodology>

<https://www.etherchain.org/account/0xea674fdde714fd979de3edf0f56aa9716b898ec8#uncles>

<https://www.kaggle.com/amritpal333/crypto-mining-data>

<https://www.kaggle.com/c/g-research-crypto-forecasting>

### Another idea

What if everybody got a solar panel at home?

* Panels on the house, north, south, east or west facing
* Predict the efficiency of how many houses you can add solar panels on

Solar flat, daily, yearly average, no need to use machine learning, since weather data we already have, and power production is set

<https://www.nature.com/articles/s41597-020-00739-0>

https://www.seia.org/solar-industry-research-data

## Wisam

Collect customer preferences, customer behaviours are harder to measure. We need purchase data, what they bought and related items they bought.

What if we were to tie it into crypto, if we can buy certain currency and see related crypto

Pick one variable and correlate it to the price of Crypto, vs time series analysis, too many methods to go into, this will hyperfocus on small trends.

Other resources wordbank, dataset.

### **Another idea (Main Idea so far)**

**Predict how many people go into retirement, see how many people we would need in the industry. We did this in module, can also try to tie into immigration. Which sectors might need to outsource labour force. Aging population is a global issue, prototype on Canada, would help in long term planning, and short term planning 5 years, how many people should replace them.**

**What data are we predicting, supervised or unsupervised, this is what canada is facing, this is the amount of people we need to replace the aging workforce. People entering university, directed towards special fields. Expats/Immigrants, I think predicting the demand, unsupervised will work better, clustering first will help us decide, that analytics field/data science wont be in danger of this issue. Farming might be. Kimchi clay pots making, running out people. Canada has census, age and occupation. Age, Occupation, Location,**

**https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/index-eng.cfm**

## Olesya

https://www.kaggle.com/imakash3011/customer-personality-analysis

## Esther

Human Behaviour and how it affects the climate, potential damage

For example:

* Restaurants and how human consumption of meat affects climate

Retail, amount of clothes people by on the daily, look at brands, x amount of dollars spent on luxury clothes vs behind the scenes the costs of a tshirt.

Look at metrics,

# Meeting meetings

* Human Behaviors bad for Machine learning
* Start with DataSet and see where you can go from there

.