Please all add to this document so that we can create our project abstract:

Mitch draft 1:

With the adoption of electric vehicles showing no signs of slowing down, the dependency on lithium ion batteries is ever increasing. This paper attempts to predict the stock price of lithium ion battery manufacturers by the stock price of lithium mining companies. The data set used includes historical financial data of the manufacturer/mining companies. We also employ sentiment analysis of reputable financial news outlets' headlines and press releases from reputable media outlets related to lithium mining. As this is a time series predictive problem and Box–Jenkins [1] suggests Autoregressive Integrated Moving Average (ARIMA) is fruitful when working in this area, ARIMA is used, as well as other machine learning algorithms to increase the accuracy of the model.

[1] G.P. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, Holden-day Inc., San Francisco, CA, 1976.

What else can we add/change to the above?

Abstract Sections:

- Project Statement:
    - Prediction of stock prices of Lithium Ion Battery manufacturers using historical data of Lithium Mining Companies.
- Brief problem description and analytics hypotheses:
    - We plan on analyzing the historical data of Lithium Mining companies for prediction. This analysis will include the stock prices of the mining companies and sentiment analysis of reputable financial news outlets' headlines regarding the same. Research of this kind could be beneficial to companies that make heavy use of lithium ion batteries by determining upcoming cost spikes by analysis of the state of the lithium mining companies.
    - Hypotheses:
        - The stock price of Li-Ion battery manufacturers is strongly correlated with the stock price of Li mining companies. (This is because Li is an important raw material for the production of batteries.)

        - The stock price movement of Li-Ion battery manufacturers is correlated to the sentiment of the relevant news articles.
- Data sources:
    - Yahoo/Google Finance
    - News articles from Bloomberg, Wall Street Journal, Financial Times, CNBC
    - Potentially publicly available weather data for mining regions
    - Potentially satellite image data of data mining regions
- Tentative timeline (using CRISP-DM Methodology):
    - Phase 1 (Business Understanding)
        - Completed
    - Phase 2 (Data Understanding) (10 days : 6th July) (**Week 1-2**)
        - Data collection (6 days: 2nd July)
        - Describe and explore data (2 days: 4th July)
        - Verify quality (2 days : 6th July)
    - Phase 3 (Data Preparation) (20 days: 26th July) (**Week 2,3,4**)
        - Data cleaning
        - Data formatting
        - Data integration
    - Phase 4 (Modeling) (10 days: 5th Aug) (**Week 4-5**)
        - E
    - Phase 5 (Evaluation and Final presentation, report) (5 days: 15th August) (**Week 6**)
        -
    - Phase 6 (Deployment) (Complete by 15th Aug)
        - There are currently no plans to deploy the model and use it exclusively for personal financial gain.
- Team members and assigned tasks
    - Mitchell
    - Anav
    - Sanchit
    - Shubham

Roadmap:
Week 1:

Goal: Data understanding
Steps:

1. Gather the data
2.

Week 1:

Goal:

Steps:

Weeks 3-4:

Goal:

Steps

Weeks 5-6:

Goal:

Steps:

: