Name: **Daniel Nazareth**
UbitName: dnazaret
Email: dnazaret@buffalo.edu

SUMMARY: The project aims to generate a time series forecast for a set of stocks using various forecasting models like the Arima model, Holt-Winters model and linear regression. Specifically the R programming language (popular for its use in statistical analysis) is used to analyse a dataset containing stock data and identify the stocks with the minimum MAE(Mean Absolute Error) which is a common metric for evaluating errors in time series analysis using the different models. Finally graphs are plotted of the 10 minimum MAE stocks in each of the models.

RESULTS: The results are summarised in tabular form below for each of the models. This can be verified by running the attached code at the command line as well.

LINEAR REGRESSION MODEL

RANK(Low to High)	STOCK	MAE VALUE
1	GRVY	0.1130815
2	ELON	0.1708333
3	SMSI	0.2753055
4	PCTI	0.3298646
5	BYFC	0.3330336
6	LFVN	0.3604605
7	TAIT	0.3694651
8	STB	0.3798104
9	STLY	0.3906403
10	BLRX	0.4163314

HOLT WINTERS MODEL

RANK(Low to High)	STOCK	MAE VALUE
1	EDS	0.06022709
2	VLYWW	0.09000000
3	IKAN	0.09451631
4	JOEZ	0.09452480
5	APWC	0.09639256
6	MTSL	0.11008672
7	COCO	0.11565898
8	HNSN	0.12703413
9	TINY	0.13458633
10	IBCA	0.13481835

ARIMA MODEL

RANK(Low to High)	STOCK	MAE VALUE
1	COCO	0.1130815
2	APWC	0.1708333
3	FREE	0.2753055
4	IKAN	0.3298646
5	SPU	0.3330336
6	ELON	0.3604605
7	VLYWW	0.3694651
8	MFI	0.3798104
9	ENZN	0.3906403
10	MTSL	0.4163314

GRAPHICAL PLOTS: Below are the results represented graphically with rank on x axis and MAE value on y for each of the three models.
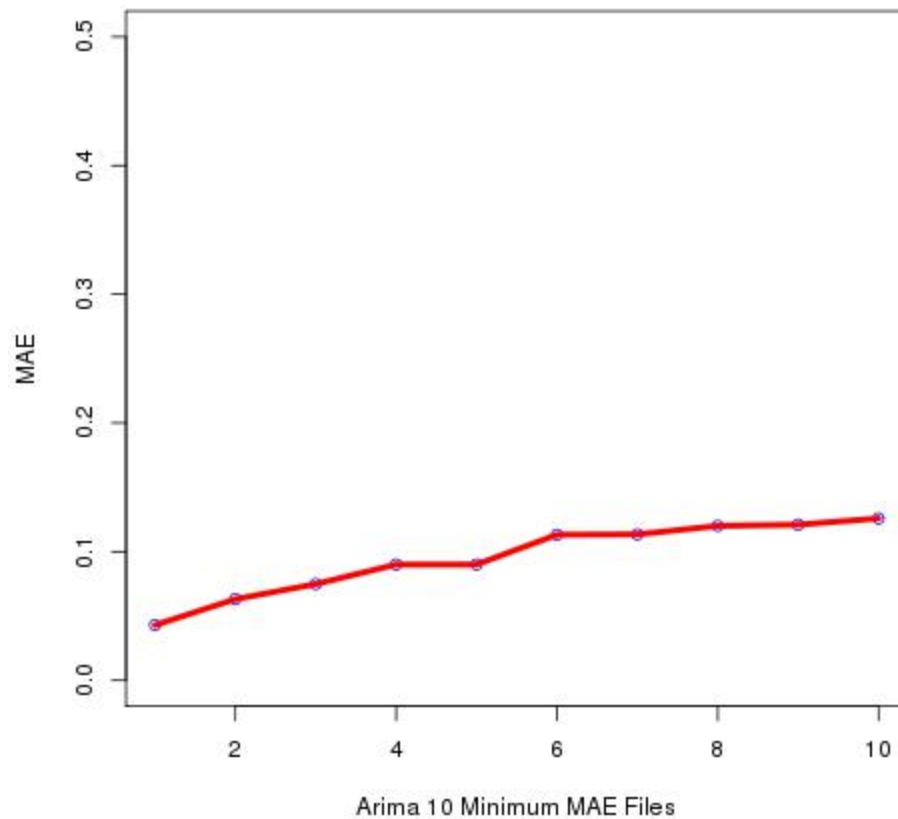
LINEAR REGRESSION MODEL

Linear Regression 10 Minimum Files

HOLT WINTERS MODEL

Holt-Winters 10 Minimum Files

ARIMA MODEL

Arima 10 Minimum MAE Files
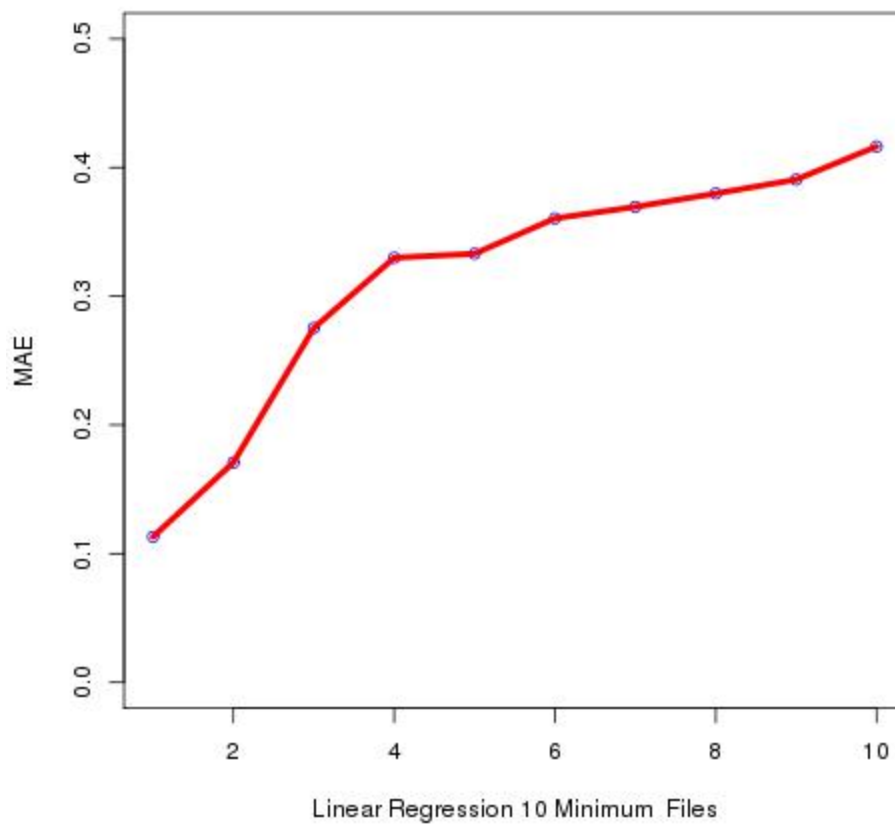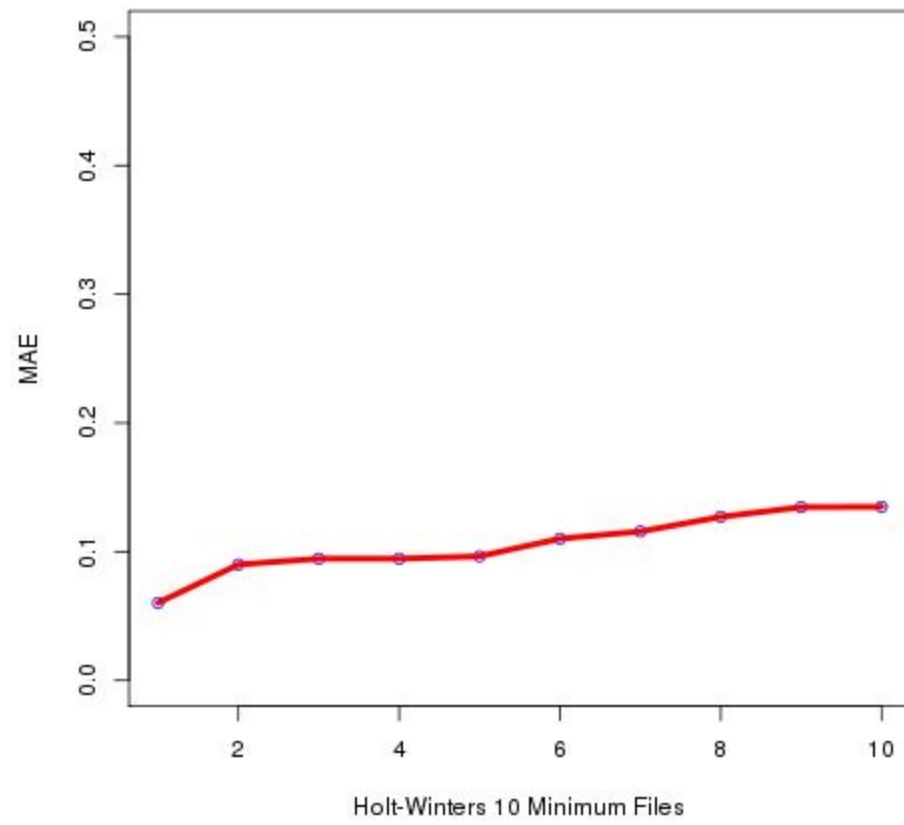
NOTES/ISSUES FACED:

- The biggest issue was the time taken to execute the ARIMA model. Even 10 hours was not sufficient to run the arima forecast on CCR. Hence due to time running out and difficulty in obtaining CCR resources for such a large amount of time, I varied the ARIMA parameters slightly (after reading the manual on the function) to reduce running time.
- Errors were encountered while running the Holt Winters model but were fixed by setting gamma=FALSE, which is a common practice while using this function.
- Some of the data files had incomplete last lines leading to warnings while running the code. However this did not pose a serious problem.
- **2 data files JASO.csv and ULTA.csv were mentioned in stocklist as JASOLtd. and ULTAInc. which led to mismatches while reading them. Hence they were edited in the code to match.**

REFERENCES

http://manuals.bioinformatics.ucr.edu/home/programming-in-r

https://stat.ethz.ch/R-manual/R-devel/library/graphics/html/plot.dataframe.html

http://www.inside-r.org/packages/cran/forecast/docs/auto.arima

http://ww2.coastal.edu/kingw/statistics/R-tutorials/simplelinear.html

https://stat.ethz.ch/R-manual/R-patched/library/stats/html/HoltWinters.html

http://a-little-book-of-r-for-time-series.readthedocs.org/en/latest/src/timeseries.html