FORECASTING RENEWABLE ENERGY USING A TIME SERIES MODEL: PREDICTING SUPPLY LEVELS IN NEW YORK STATE

Group 5:
Mike Barbiere, Chigozie Nnodim, Estherline Pierre, Cassidy Zaruba

Table of Contents

Organization & Problem Context

- Limited amount of green-energy supply available, and the amount of energy consumption being experienced is currently high
- Forecast future levels of green-energy supply and energy consumption
- When forecasting future levels of green-energy supply and energy consumption, businesses will need to understand the current situation of the problem.
- The New York area currently has a limited amount of green-energy supply available.
- Forecasting future levels of green-energy supply and energy consumption allows businesses to plan for the necessary infrastructure changes and ensure that there is enough green energy available to meet the needs of the population.

Solv4x: The Business Problem

- As Mass adoption of Electric Vehicles (EV) becomes increasingly popular and more widespread, the demand for energy will continue to grow.
- Businesses can increase the supply of green energy by increasing the number of renewable energy sources, such as solar and wind power
- Solv4x has been examining the proportion of energy-load demand that’s supplied by renewables using Machine Learning (ML) forecasting algorithms.
- Our project’s main focus is to predict the trend of renewable energy supply as a proportion of overall load-demand.

Analytical Problem

Part 1

- Identify the data source for renewable energy and daily load-demand.

Part 2

- Construct forecasting model to determine when energy demand will increase with renewable generation.

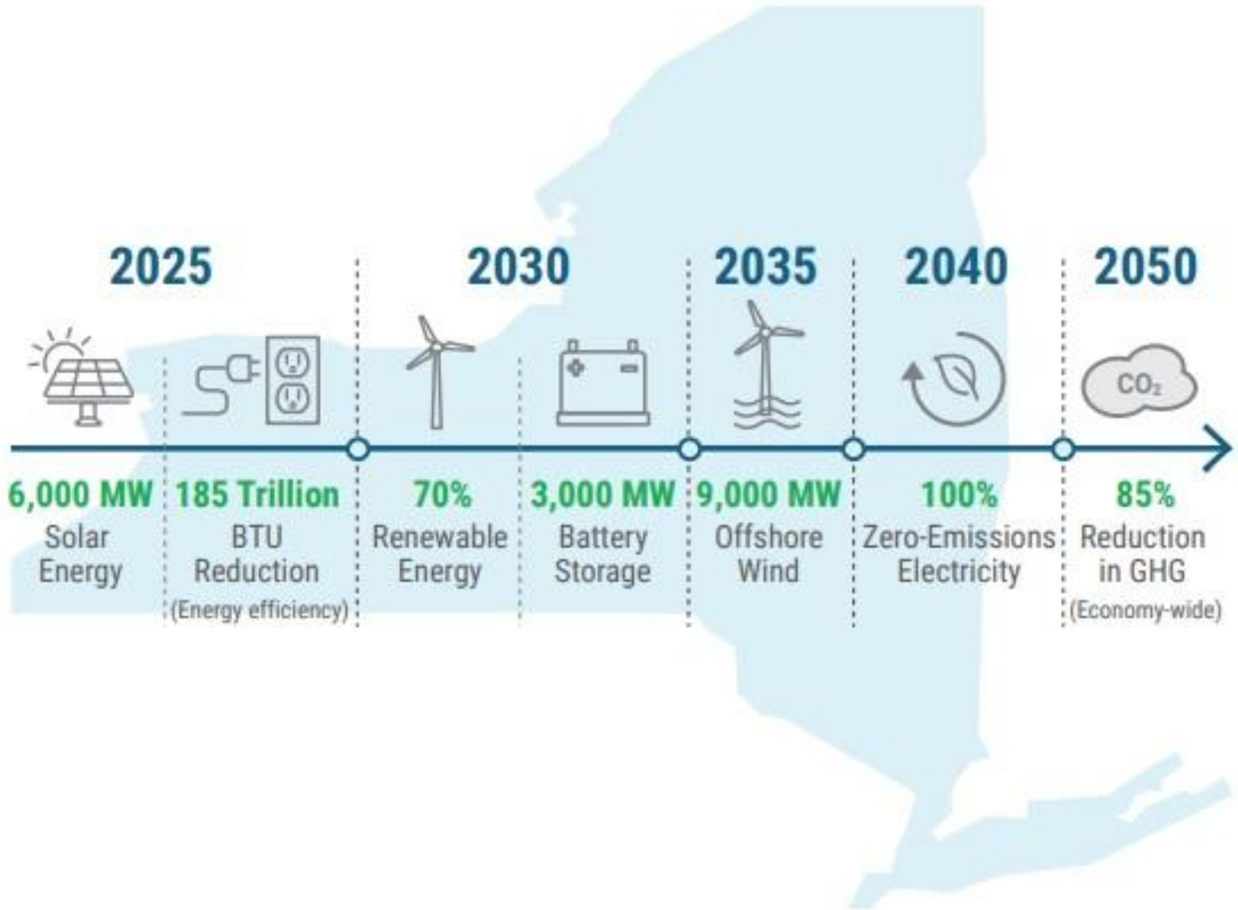
Part 3

- Create a report comparing the forecast and realized values.

Part 1: Identify Data Source

Searching and researching.

Part 1: New York Independent System Operator (ISO)



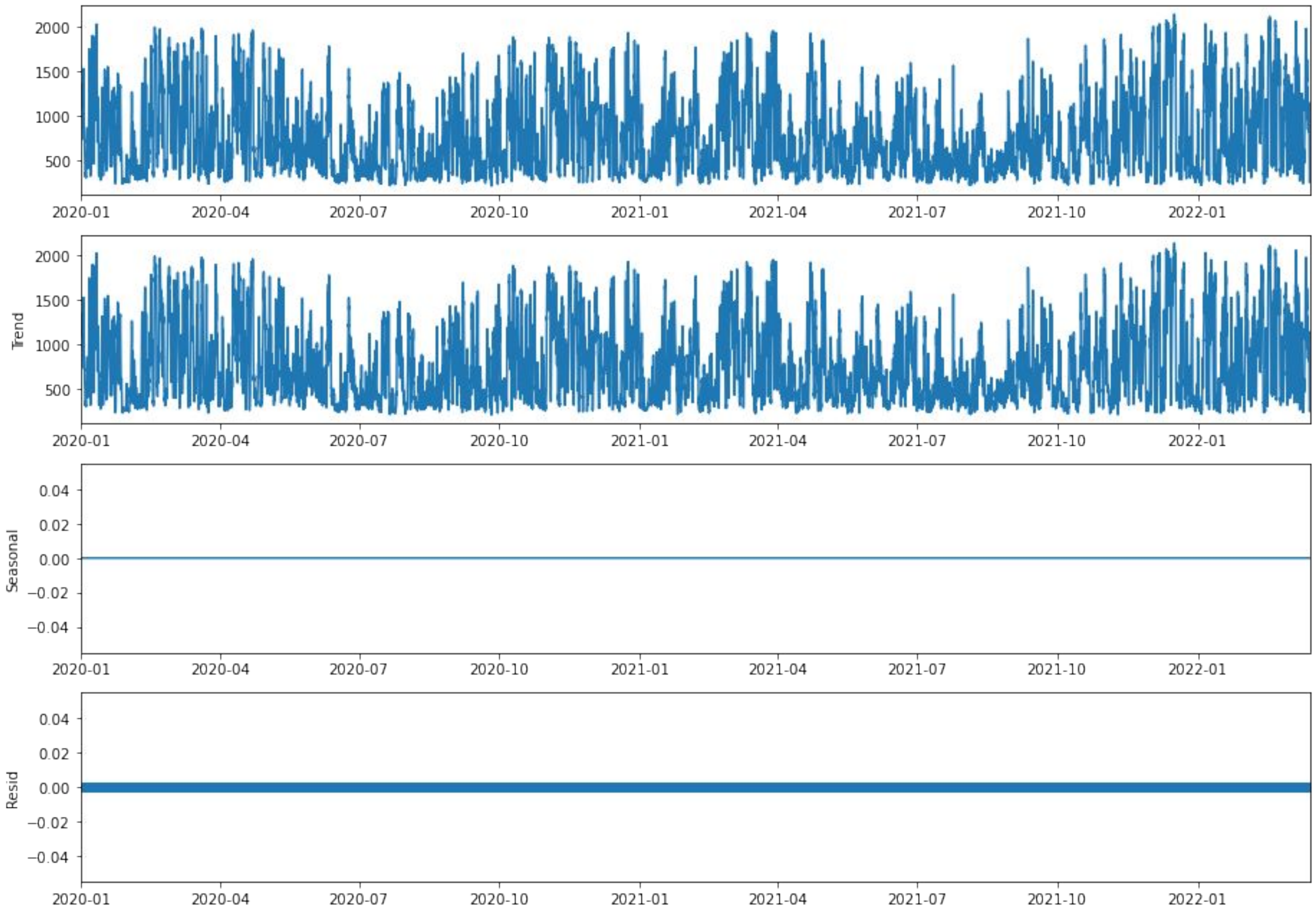
- Total energy load data (i.e., demand) with forecasted values
- Proportion of energy load supplied by Fuel Category type, including renewable energy sources
- *Business Problem Processes*:
  - Reach out to the NY ISO to determine if data are available
  - Construct datasets using historical values.

Part 2: Constructing Forecasting Model

Development and implementation.

Our Dataset

- Dataset construction: time-consuming, cumbersome in Excel
- Data preprocessing: easily mutable, computationally inexpensive.
- Model construction: learning how to use the ARIMA model



ARIMA Model Fitting

Time Stamp	Gen MW	Forecast	Difference
2022-03-16 00:00:00	361.90	456.76	0.792
2022-03-16 01:00:00	371.16	476.04	0.779
2022-03-16 02:00:00	381.00	494.42	0.770
2022-03-16 03:00:00	404.66	511.84	0.790
2022-03-16 04:00:00	446.41	528.27	0.845

Mean Difference: 0.696

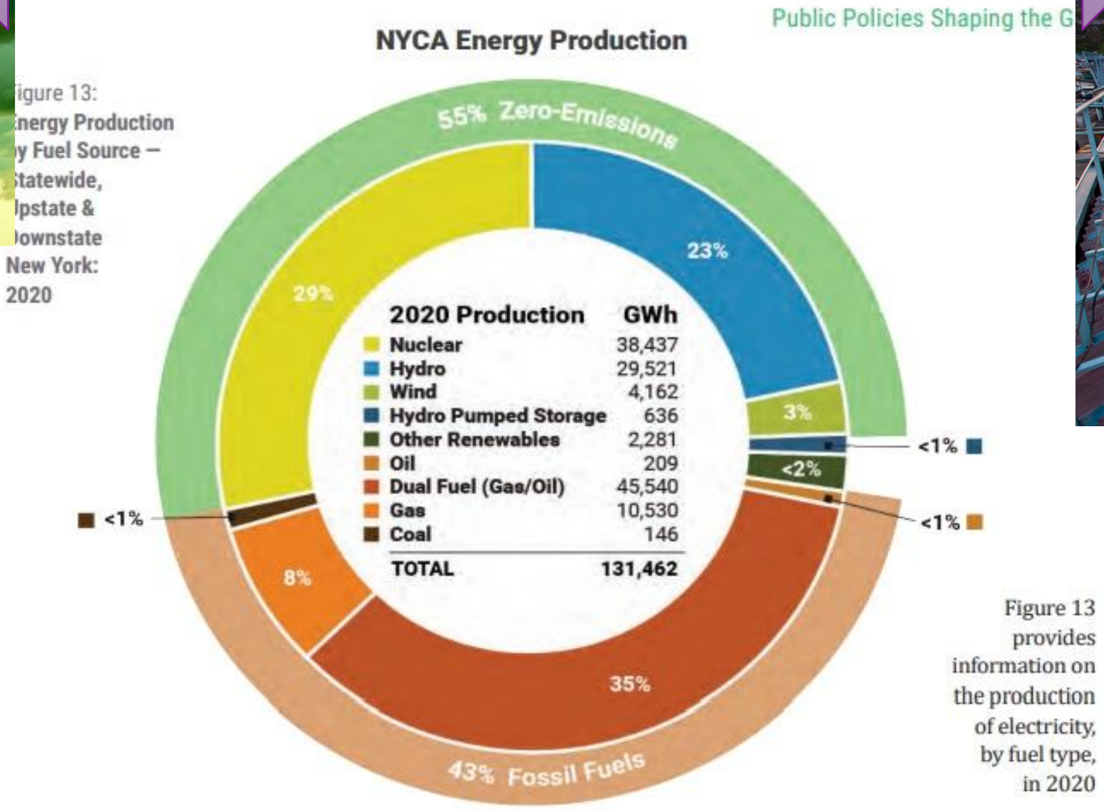
```
Performing stepwise search to minimize aic
 ARIMA(1,0,1)(0,0,0)[0]             : AIC=1955253.833, Time=19.84 sec
 ARIMA(0,0,0)(0,0,0)[0]             : AIC=3671244.505, Time=2.95 sec
 ARIMA(1,0,0)(0,0,0)[0]             : AIC=inf, Time=3.76 sec
 ARIMA(0,0,1)(0,0,0)[0]             : AIC=3364827.445, Time=39.63 sec
 ARIMA(2,0,1)(0,0,0)[0]             : AIC=1954477.868, Time=25.63 sec
 ARIMA(2,0,0)(0,0,0)[0]             : AIC=inf, Time=4.91 sec
 ARIMA(3,0,1)(0,0,0)[0]             : AIC=1954623.422, Time=148.81 sec
 ARIMA(2,0,2)(0,0,0)[0]             : AIC=1952760.272, Time=127.57 sec
 ARIMA(1,0,2)(0,0,0)[0]             : AIC=1955005.834, Time=30.56 sec
 ARIMA(3,0,2)(0,0,0)[0]             : AIC=1953181.642, Time=93.53 sec
 ARIMA(2,0,3)(0,0,0)[0]             : AIC=1953538.382, Time=63.85 sec
 ARIMA(1,0,3)(0,0,0)[0]             : AIC=1953655.235, Time=19.14 sec
 ARIMA(3,0,3)(0,0,0)[0]             : AIC=1952854.646, Time=118.79 sec
 ARIMA(2,0,2)(0,0,0)[0] intercept   : AIC=1952381.604, Time=387.16 sec
 ARIMA(1,0,2)(0,0,0)[0] intercept   : AIC=1954883.085, Time=124.50 sec
 ARIMA(2,0,1)(0,0,0)[0] intercept   : AIC=1954349.104, Time=67.18 sec
 ARIMA(3,0,2)(0,0,0)[0] intercept   : AIC=1952391.244, Time=476.92 sec
 ARIMA(2,0,3)(0,0,0)[0] intercept   : AIC=1953493.044, Time=187.23 sec
 ARIMA(1,0,1)(0,0,0)[0] intercept   : AIC=1955136.940, Time=56.66 sec
 ARIMA(1,0,3)(0,0,0)[0] intercept   : AIC=1953518.515, Time=142.99 sec
 ARIMA(3,0,1)(0,0,0)[0] intercept   : AIC=1954442.087, Time=221.44 sec
 ARIMA(3,0,3)(0,0,0)[0] intercept   : AIC=1952790.392, Time=589.98 sec

Best model:  ARIMA(2,0,2)(0,0,0)[0] intercept
Total fit time: 2956.094 seconds
```

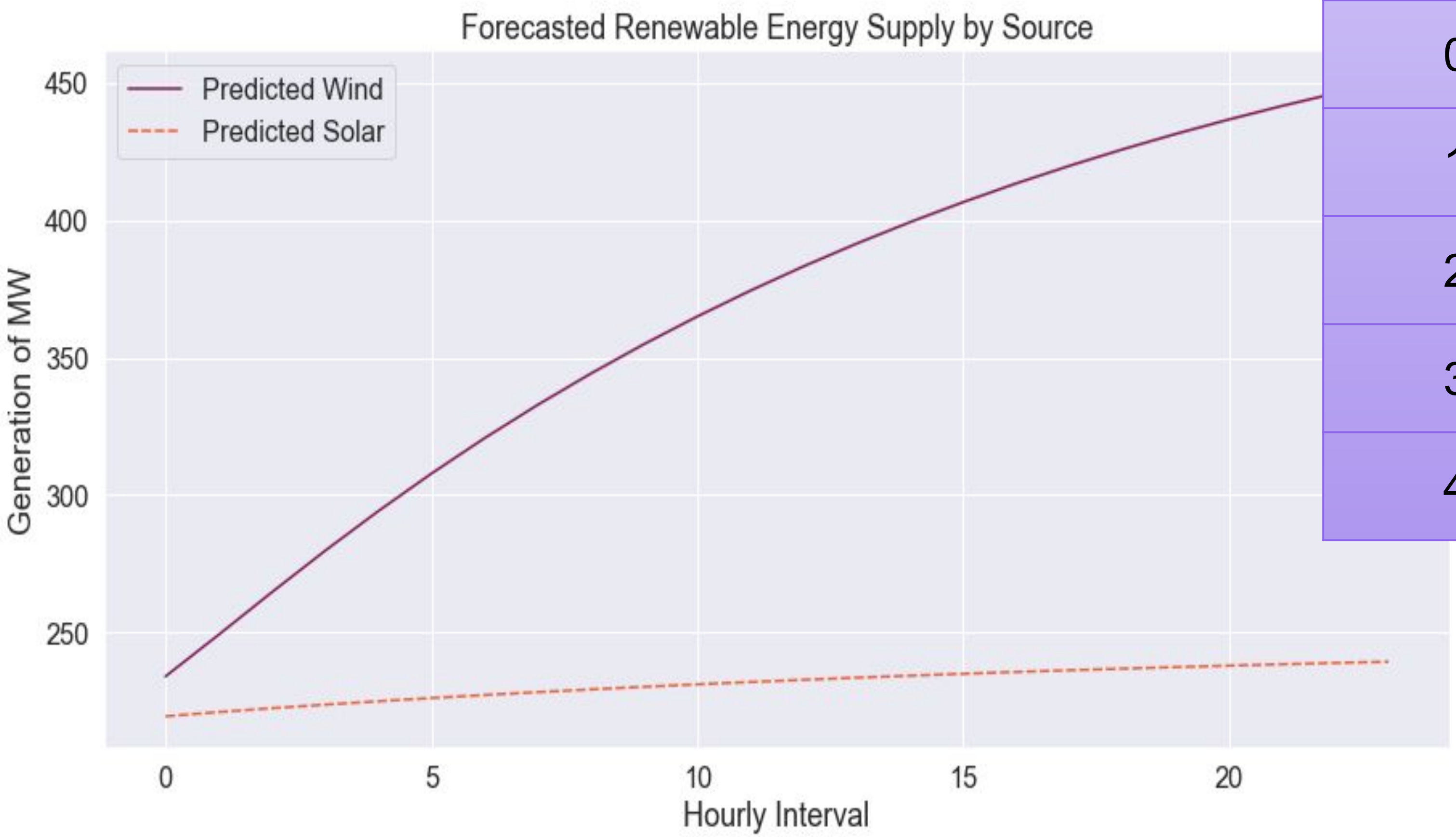
Part 2 Cont'd: Multivariate Model Construction

Disaggregating solar and wind energy.

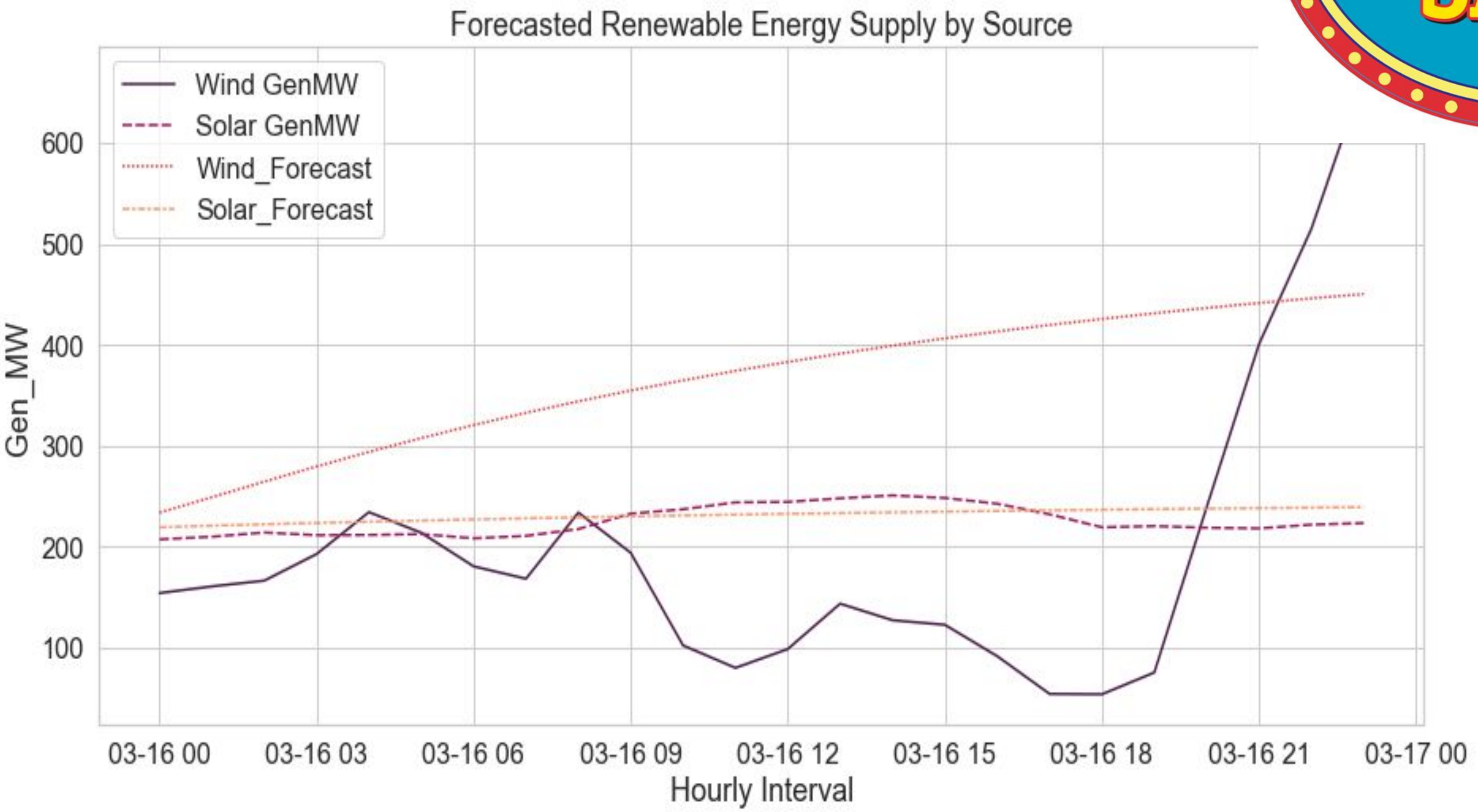
Renewable Energy Supply



VARMAX Multivariate Forecasting Model



Hour-Interval	Predict_Wind	Predict_Solar
0	234.036392	219.536305
1	249.189231	221.082089
2	264.637814	222.476174
3	279.707199	223.778195
4	294.122424	225.007667



Mean Wind: 0.538

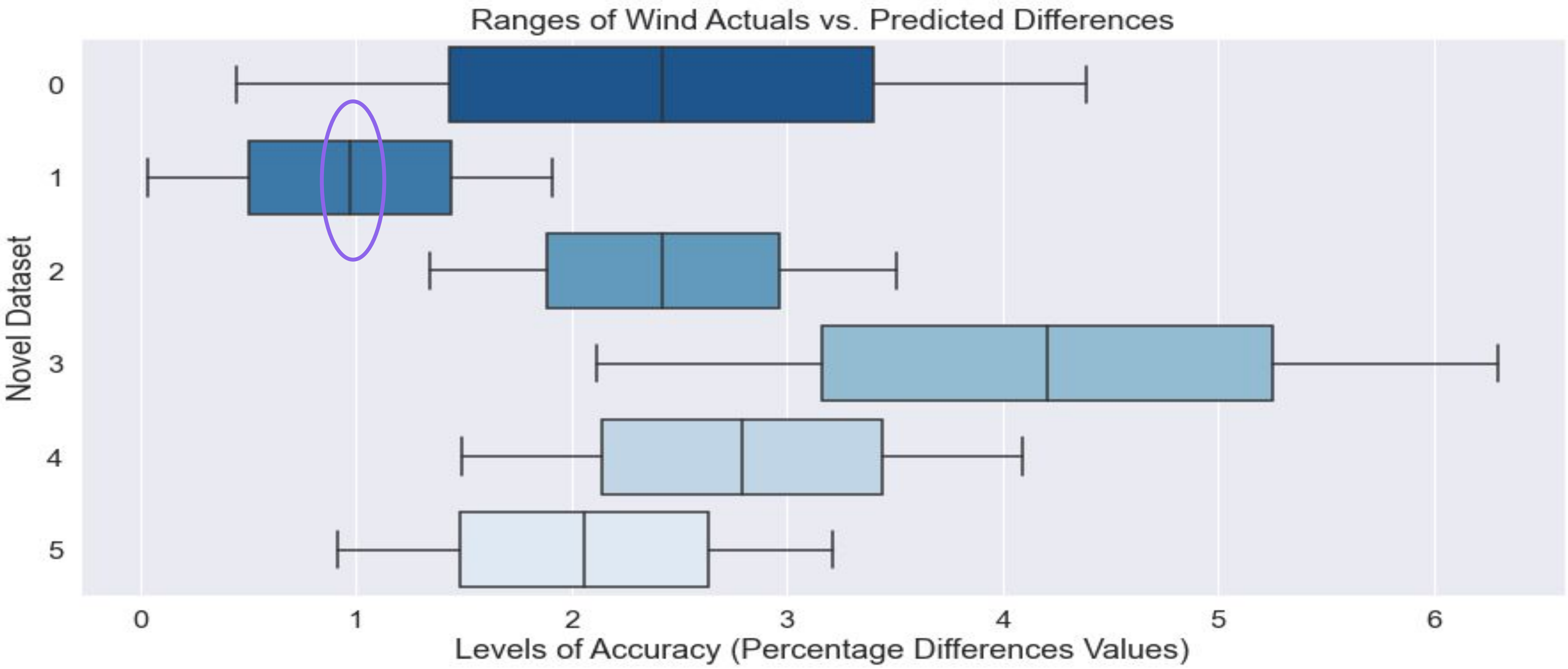
Mean Solar: 0.974

Total Avg: 0.756

Analysis of Accuracy

Time Stamp	Wind GenMW	Solar GenMW	Wind_Forecast	Solar_Forecast	Diff_Wind	Diff_Solar
2022-03-16 00:00:00	154.363636	207.545455	234.036392	219.536305	0.659571	0.945381
2022-03-16 01:00:00	161.000000	210.166667	249.189231	221.082089	0.646095	0.950627
2022-03-16 02:00:00	166.666667	214.333333	264.637814	222.476174	0.629792	0.963399
2022-03-16 03:00:00	193.000000	211.666667	279.707199	223.778195	0.690007	0.945877
2022-03-16 04:00:00	234.500000	211.916667	294.122424	225.007667	0.797287	0.941820

Variability of Wind Accuracy



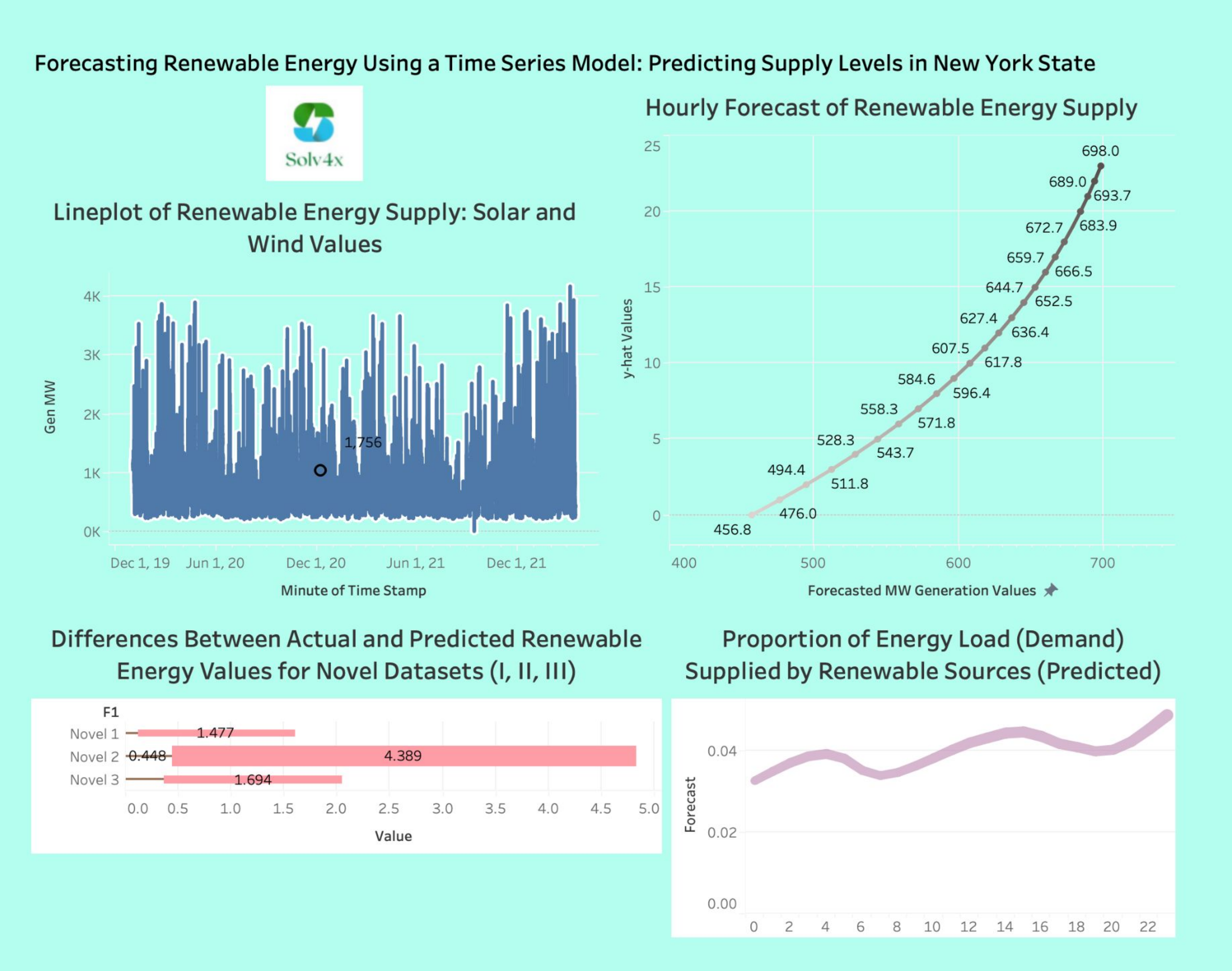
Part 3: Creating A Report

What have we done?

Findings & Results

- When using the VARMAX model to forecast wind values, our resulting RMSE was 220.98, which is fairly large given the average Wind-energy value as reported by NYISO for all days in 2020-21 as 482.14.
  - Ultimately, our model's error-margin was nearly half the value of any given wind-supply data (220.98 / 482.14 = 0.46).
- Although the RMSE and calculated accuracy levels were unsatisfactory; Solv4x expressed their satisfaction in our reporting, as our group was only one in a larger network of teams.
  - Our Business Sponsor Rekha expressed her intentions in using the foundational findings to construct a more complex forecasting model that would account for numerous other features (e.g., other potential input variables to enhance predictive power).
  - It's possible that Solv4x may combine our results with an LSTM or ANN model that was developed elsewhere.
- On the other hand, our Solar forecasting values were highly accurate, and using the ARIMA model may be an acceptable forecasting solution for states with an increased emphasis on this energy source.
  - For example, California absorbed 15.7%, or 30 TWh, of its energy needs in 2020 from solar sources (Jacobs, 2021); compared to New York's solar energy values, which was approximately 2.5% of supply (wind comprised 4% and hydroelectric ~12%) (US EIA, n.d.).
  - Therefore, specific geographic considerations should also be considered when potentially implementing the model.

Extensions of Our Work: Tableau Dashboard



CONCLUSION

- If Group 5 was allotted more time, we would continue our research, and apply black-box models or LSTM model to see if they provided a better accuracy rating.

- Group 5 is resolving Solv4x's business problem through our coding, outside research, hands-on learning, and visualizations. Our assignment provided a significant perspective for Solv4X, and the company will continue its assignment.

Sources

Alsharif, M. H., Younes, M. K., & Kim, J. (2019). Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea. Symmetry, 11(2), 240.

Lai, J. P., Chang, Y. M., Chen, C. H., & Pai, P. F. (2020). A survey of machine learning models in renewable energy predictions. *Applied Sciences*, *10*(17), 5975.

Foley, A. M., Leahy, P. G., Marvuglia, A., & McKeogh, E. J. (2012). Current methods and advances in forecasting of wind power generation. Renewable energy, 37(1), 1-8. doi:10.1016/j.renene.2011.05.033

Jacobs, F.. (2021, September 28). *Electricity Generation by US State.* Big Think. Accessed April 22, 2022 from Electricity generation by U.S. state - Big Think Energy mix per U.S. state

NYC.gov. (n.d.). *NYC is building a clean, resilient, and affordable energy system.* Accessed February 10, 2022, from Energy - Sustainability (nyc.gov)

NYISO (n.d.). *Power Trends 2021, Annual Grid & Markets Report.* Accessed February 14, 2022, from Power Trends - NYISO

Perera, K. S., Aung, Z., & Woon, W. L. (2014, September). Machine learning techniques for supporting renewable energy generation and integration: a survey. In *International Workshop on Data Analytics for Renewable Energy Integration* (pp. 81-96). Springer, Cham.

Sfetsos, A. (2000). A comparison of various forecasting techniques applied to mean hourly wind speed time series. Renewable energy, 21(1), 23-35.

U.S. Energy Information Administration. (n.d.). New York, State Profile and Energy Estimates. www.eia.gov, Accessed April 22, 2022, from https://www.eia.gov/state/analysis.php?sid=NY

Thank You