DATA BOSSES

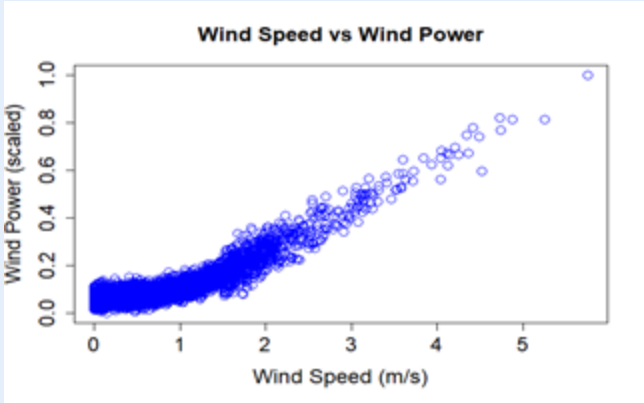
Christine Domercant (acd201)
Truc Hoang (tth46)
Tasnim Quayum (tnq3)

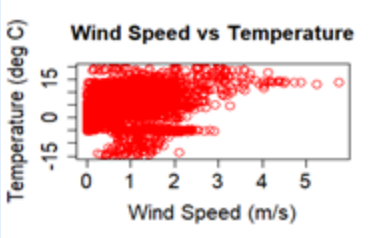
Descriptive Analytics

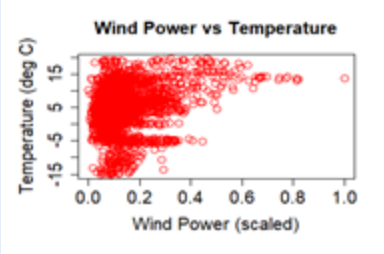
Analysis

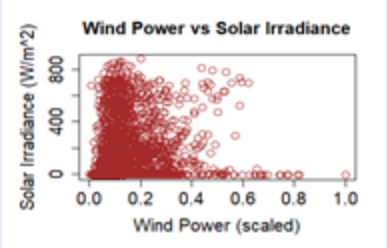
Visualization

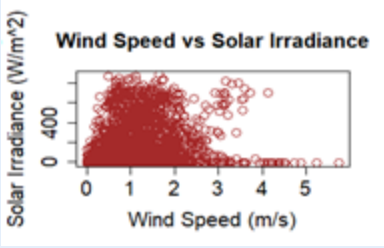
Created various plots to visually analyze the relationship between the variables











Correlation

Computed the correlation to strength visual analysis

```
> cor(data_new)
                  Speed_ms Temperature_C     Humidity Solar_Irradiance_Wm2 Power_scaled
Speed_ms         1.0000000   0.274028779 -0.169861218            0.2765546    0.8885186
Temperature_C    0.2740288   1.000000000 -0.004808583            0.2133071    0.2570446
Humidity        -0.1698612  -0.004808583  1.000000000           -0.4658484   -0.0675234
Solar_Irradiance_Wm2 0.2765546 0.213307149 -0.465848366         1.0000000    0.1958399
Power_scaled     0.8885186   0.257044591 -0.067523396            0.1958399    1.0000000
```

Pseudocode

Brainstormed how parts mentioned above can be used to implement prediction

Wind Power Prediction Modeling Efforts

Linear

High correlation between wind power and wind speed

Exponential

Scatter plot showed slight curvature

Polynomial

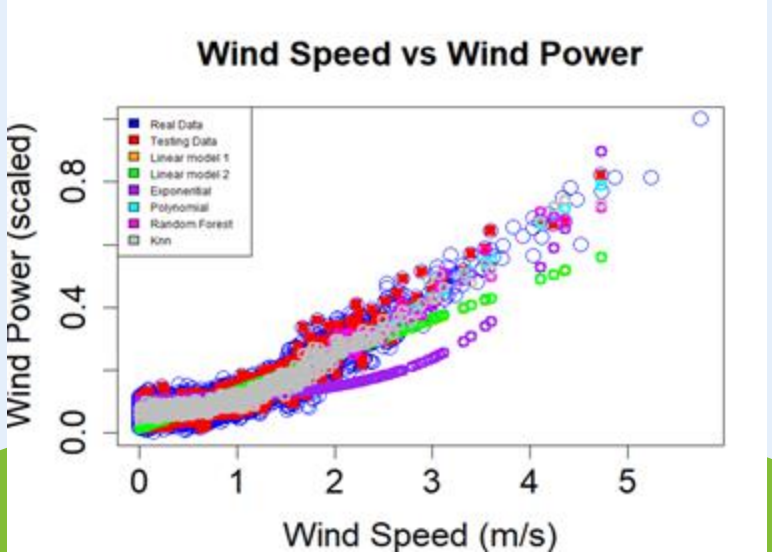
Capture the nonlinear relationship

Random Forest

Model the complex relationship in prediction

Knn

Investigate non parameter approach



Regression Model	Train_RMSE	Test_RMSE
Linear	0.042	0.042
Exponential	0.064	0.073
Polynomial	0.029	0.029
Random Forest	0.015	0.031
Knn	0.025	0.032

Wind Speed Forecasting Modeling Efforts

01 ARIMA

Use historical wind speed data to forecast - assume time series data to be stationary (removing trend and seasonality)

02 ARFIMA

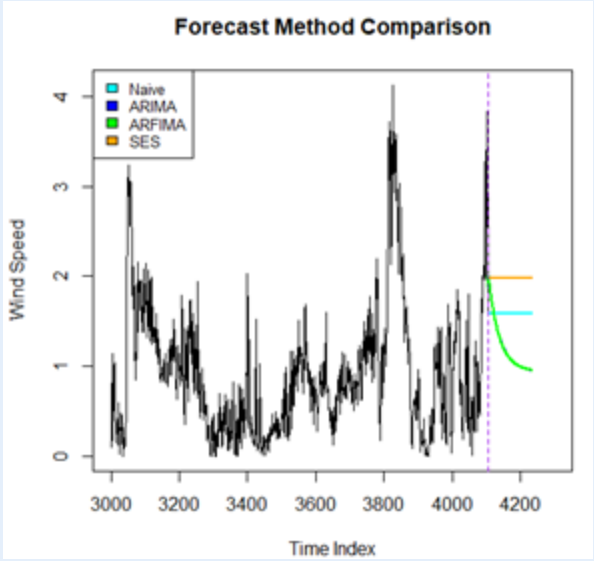
Use historical wind speed data to forecast - handle non-stationary data by introducing a fractionally integrated component.

03 SES

Use a weighted average of past observations to forecast future values

04 ARIMAX

Extend ARIMA by allowing the inclusion of exogenous variables in the model (Temperature and Solar Irradiance)



Final Model Description

ARFIMA

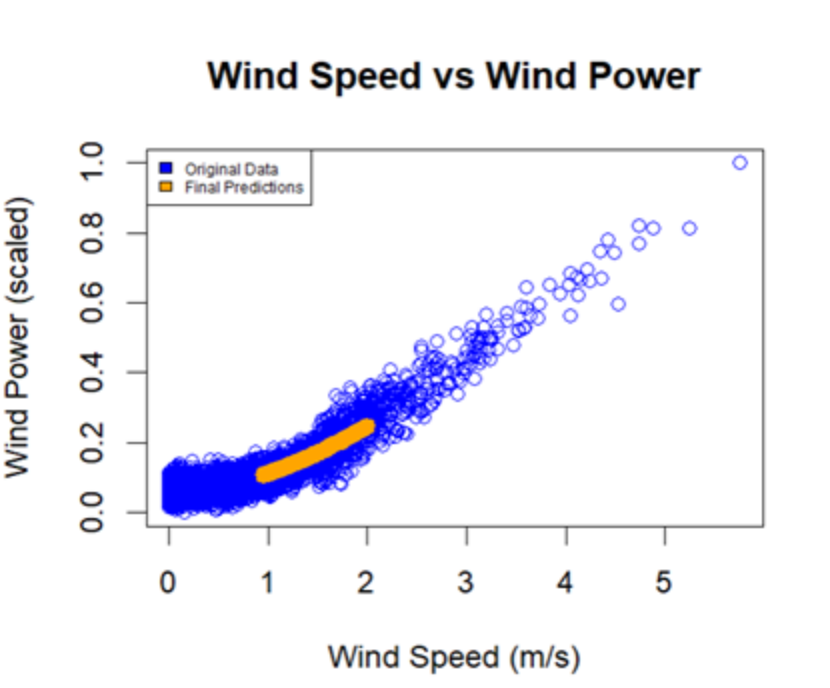
Polynomial

```
Power_scaled ~ Temperature_C + I(Speed_ms^2) + I(Speed_ms^3)
```

Forecast Model	RMSE Wind speed	RMSE PowerScale (Linear Regression)	RMSE PowerScale (Polynomial Regression)
Naive	0.6380118	0	0
ARIMA	0.5599901	0.0596	0.0544
ARFIMA	0.5599882	0.0596	0.0544
Exponential Smoothing	0.6157661	0.0605	0.0603
ARIMAX	0.6009415	0.06	0.058

Table: RMSE of wind speed from different time series forecast models and its RMSE of Wind Scaled power by Polynomial regression model.

Results and Findings

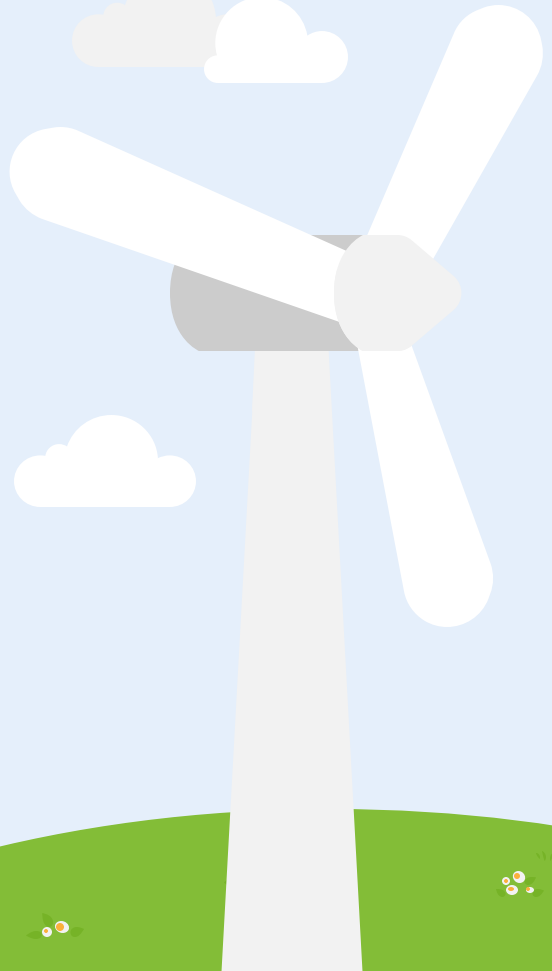


Wind speed is predicted using ARFIMA model.

Wind power is predicted using polynomial model.

All historical data was used to train models.

Recommendations & Takeaways



Recommendations:

- Continue to improve forecasting and regression model.
- Use larger data set to better understand trends over time.

Takeaways:

- There are many ways to predict future events.
- It is a continuous process.