ADS1001_GROUP2_COLAB

June 3, 2025

1 Group 2: "How do wind speed, gust patterns, and wind chill vary across different regions in Malaysia and what are their implications for the renewable energy sector?"

1.1 Group Members:

- Medellin, 34537791
- Jun Siang, 35745673
- Qi Zhi, 36222682
- Sharon, 35887311
- Anusha, 35216603
- Jia Yi, 35052708

1.1.1 Project Progress:

30 March 2025 1. Dataset Selection by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi

9 April 2025 1. Topic question formed by Anusha 2. Sub-questions formed by Anusha

16 April 2025 1. Figures drafted for sub-questions by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi

19 April 2025 1. Figures generated by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi 2. Figures analysed by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi

30 April 2025 1. Project Repository created on Github by Anusha 2. Data analyses and visualizations uploaded on project repository by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi

7 May 2025 1. Collaborative-coding platform created on Google Colab by Anusha 2. Data analyses and visualizations uploaded on Colab file by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi

14 May 2025 1. Revised figures generated by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi 2. Revised figures analysed by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi

16 May 2025 1. Fixed project structure formatting by Anusha 2. Fixed graph numbering by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi

21 May 2025 1. Revised analyses by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi

24 May 2025 1. Report conclusion drafted and revised by Anusha 2. Final revision of figures and analyses by Medellin, Jun Siang, Qi Zhi, Sharon, Anusha, Jia Yi

1.1.2 Project Background Information

As Malaysia pushes for renewable energy, understanding wind patterns is key. This project examines regional variations in wind speed, gust patterns, and wind chill across Malaysia, using historical weather data to evaluate their implications for wind energy development. Our findings aim to inform stakeholders on the suitability of different regions for renewable energy investment and planning.

1.2 Import Statements and Data Loading:

[2]:
```
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import PolynomialFeatures
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt


df = pd.read_csv("malaysia_weather_data.csv", index_col = False, sep = ',')
```

1.2.1 Research Question 1: "How do wind speed and gust intensity vary across different regions of Malaysia?"

Medellin

1.2.2 Solution:

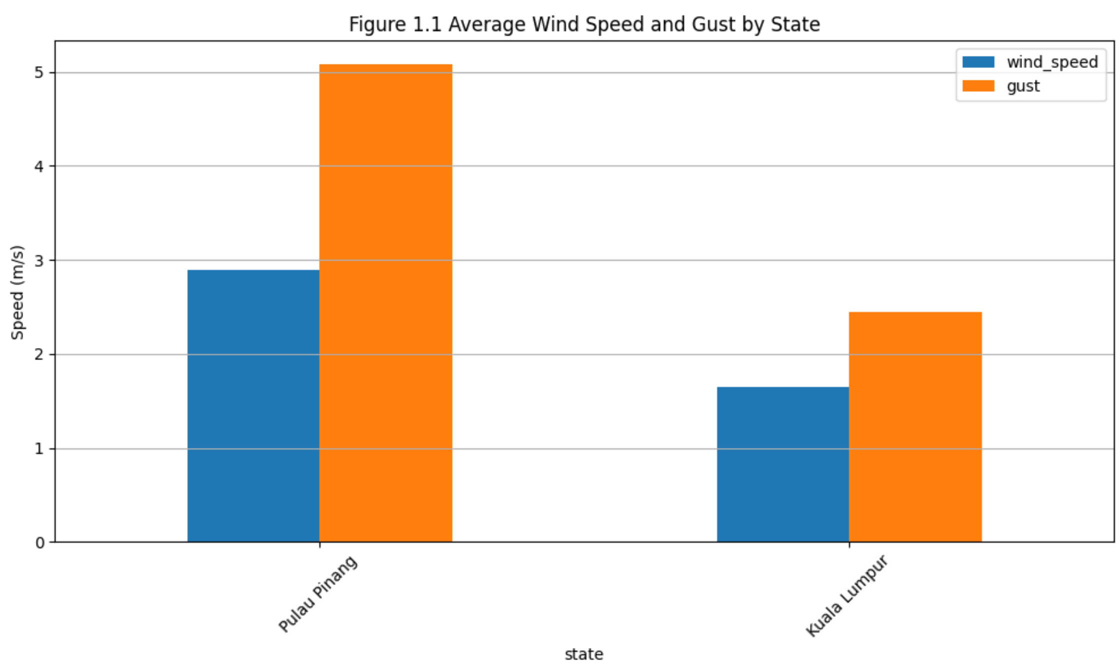
[ ]:
```
df_clean = df.dropna(subset=["wind_speed", "gust"])
df_clean.head()
```

[ ]:
```
                  place          city          state  temperature  pressure  \
0  Bandar Sri Permaisuri  Kuala Lumpur  Kuala Lumpur         30.3  1015.325
1           Ampang Jaya  Kuala Lumpur  Kuala Lumpur         27.0  1000.680
2           Bukit Jalil  Kuala Lumpur  Kuala Lumpur         26.4  1013.550
3           Bukit Jalil  Kuala Lumpur  Kuala Lumpur         31.1  1005.420
4           Penang Road   George Town  Pulau Pinang         25.8  1007.790

   dew_point  humidity  wind_speed  gust  wind_chill  uv_index  \
0       25.1      73.7         0.9   1.5        30.3       1.0
1       25.6      92.0         5.2   5.3        27.0       0.0
2       26.3      99.0         0.0   0.0        26.4       NaN
3       30.9      99.0         0.0   0.0        31.1       NaN
4       23.9      88.0         6.4  16.1        25.8       NaN

   precipitation_rate  precipitation_total  year month  day  hour  minutes  \
0                 0.0                 0.00  2022   Sep   23    13       54
```

```
1                   0.0                 17.78  2023   May    4    19         9
2                   0.0                 32.00  2021   Oct   23    23        44
3                   0.0                  0.00  2021   Mar   21    18         9
4                   0.0                  0.00  2022   Dec   22     7         0

   seconds
0       57
1       59
2       43
3       47
4       23
```

Bar chart showing average wind speed and gust per state

```
[ ]: df_clean = df.dropna(subset=["wind_speed", "gust"])

     # 1. Bar chart showing average wind speed and gust per state
     state_avg = df_clean.groupby("state")[["wind_speed", "gust"]].mean().
      ↪sort_values(by="wind_speed", ascending=False)

     state_avg.plot(kind='bar', figsize=(10, 6))
     plt.title("Figure 1.1 Average Wind Speed and Gust by State")
     plt.ylabel("Speed (m/s)")
     plt.xticks(rotation=45)
     plt.grid(axis='y')
     plt.tight_layout()
     plt.show()
```



1.2.3 Purpose:

To observe average wind behavior across different regions.

1.2.4 Insight:

Penang consistently shows higher mean wind speeds and gusts compared to Kuala Lumpur.

This indicates Penang may be more suitable for wind energy generation or applications sensitive to wind conditions.

Top Performing Places (based on average wind metrics):

Batu Maung, Sungai Puyu, Sepang

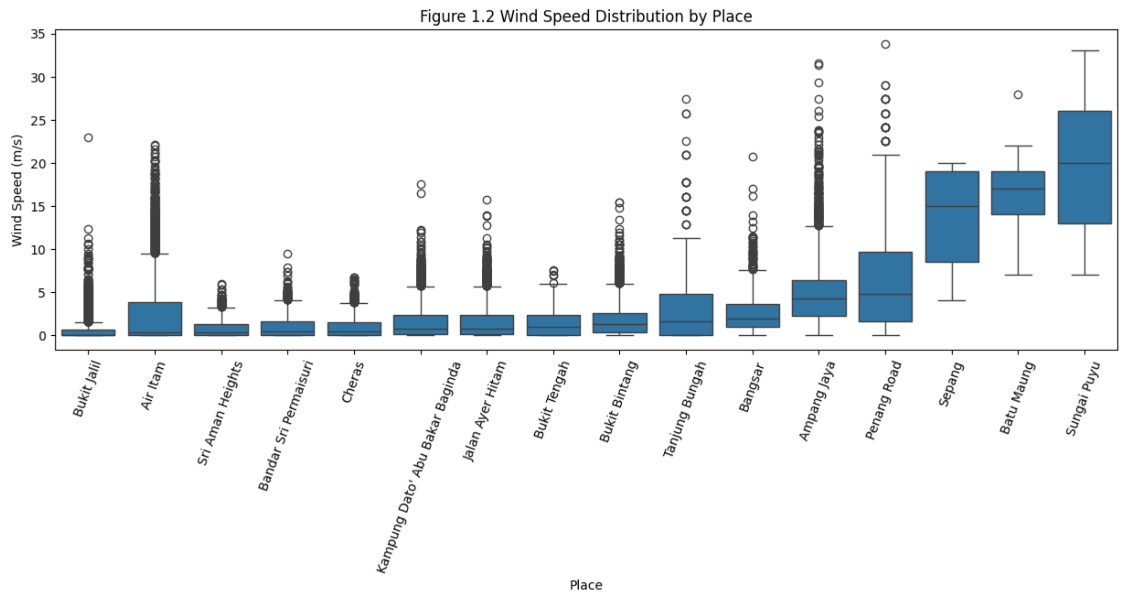
1.2.5 Implications

Energy companies have a better opportunity in Penang if wishing to venture into Wind energy sector.

Gust is higher in Penang so more precautions must also be taken.

Boxplot of wind speed and gust by place

[ ]:
```
# Sort places by median wind_speed
wind_order = df_clean.groupby('place')['wind_speed'].median().sort_values().
 ↪index

# Sort places by median gust
gust_order = df_clean.groupby('place')['gust'].median().sort_values().index

plt.figure(figsize=(15, 10))

# Sorted by wind_speed median
plt.subplot(2, 1, 1)
sns.boxplot(x="place", y="wind_speed", data=df_clean, order=wind_order)
plt.title("Figure 1.2 Wind Speed Distribution by Place")
plt.ylabel("Wind Speed (m/s)")
plt.xlabel("Place")
plt.xticks(rotation=70)

plt.show()
```



1.2.6 Purpose:

To understand the variability and spread of wind behavior, including extreme values.

1.2.7 Insight:

Distribution in Penang shows a wider range, suggesting more variability and frequent high wind events, due to variablity of coastal and inland regions.

Some locations demonstrate extreme wind conditions, with wind speed and/or gusts exceeding 25 m/s.

Sungai Puyu, Batu Maung emerges as a top candidate location with both high median, ideal for projects requiring high wind activity.

Outliers in the graph shows particular stormy weather in otherwise calm places.

1.2.8 Implications

The aformentioned places have better potential for wind farms.

Variability is also high so back up energy must still be available.

1.2.9 Research Question 2: "What is the monthly variability of wind speed and gust intensity, and how consistent are these patterns across regions?"

Jun Siang

1.2.10 Solution:

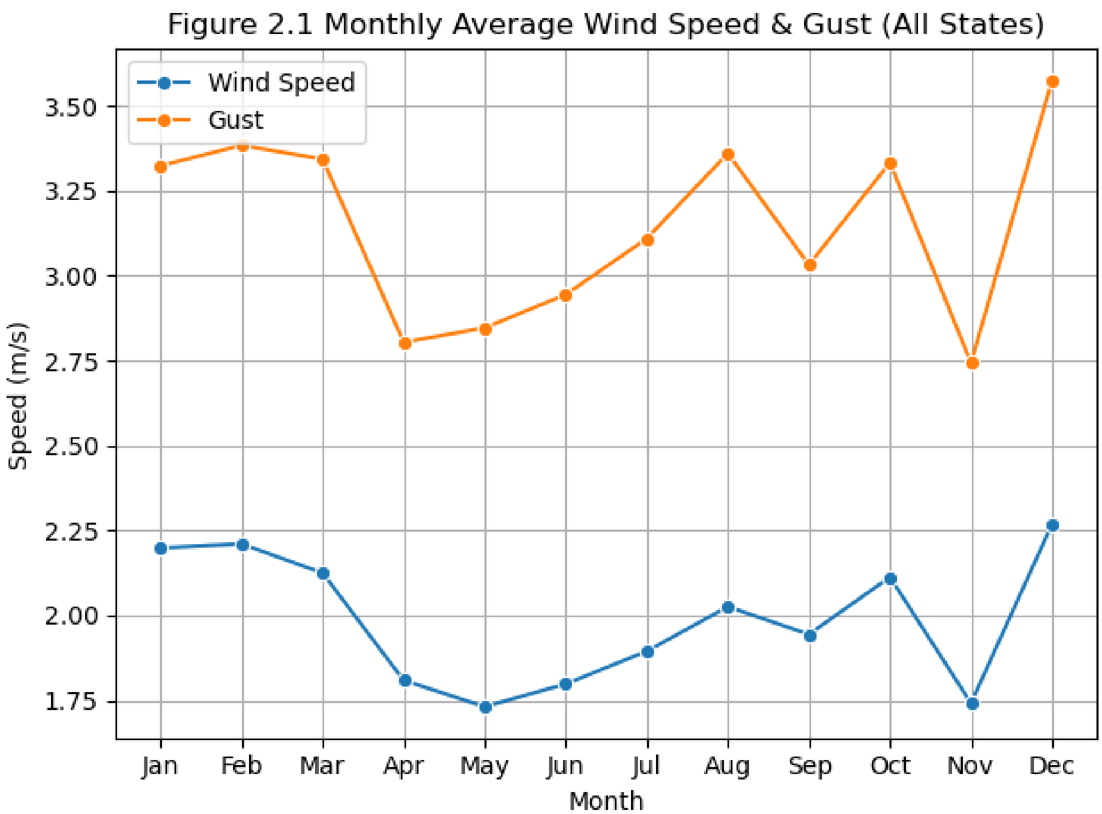
[3]:
```
month_order = ['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun',
               'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec']
df_clean2 = df.dropna(subset=["wind_speed", "gust"])
df_clean2 = df_clean2.copy()
df_clean2["month"] = pd.Categorical(df_clean2["month"], categories=month_order,␣
 ↪ordered=True)
df_clean2.head()
```

[3]:
```
                    place          city          state  temperature  pressure  \
0  Bandar Sri Permaisuri  Kuala Lumpur  Kuala Lumpur          30.3  1015.325
1            Ampang Jaya  Kuala Lumpur  Kuala Lumpur          27.0  1000.680
2            Bukit Jalil  Kuala Lumpur  Kuala Lumpur          26.4  1013.550
3            Bukit Jalil  Kuala Lumpur  Kuala Lumpur          31.1  1005.420
4            Penang Road   George Town  Pulau Pinang          25.8  1007.790

   dew_point  humidity  wind_speed  gust  wind_chill  uv_index  \
0       25.1      73.7         0.9   1.5        30.3       1.0
1       25.6      92.0         5.2   5.3        27.0       0.0
2       26.3      99.0         0.0   0.0        26.4       NaN
3       30.9      99.0         0.0   0.0        31.1       NaN
4       23.9      88.0         6.4  16.1        25.8       NaN

   precipitation_rate  precipitation_total  year month  day  hour  minutes  \
0                 0.0                 0.00  2022   Sep   23    13       54
1                 0.0                17.78  2023   May    4    19        9
2                 0.0                32.00  2021   Oct   23    23       44
3                 0.0                 0.00  2021   Mar   21    18        9
4                 0.0                 0.00  2022   Dec   22     7        0

   seconds
0       57
1       59
2       43
3       47
4       23
```

Overall average wind speed and gust by month

[4]:
```
monthly_avg = df_clean2.groupby("month", observed=True)[["wind_speed", "gust"]].
 ↪mean().reset_index()
monthly_avg["month"] = pd.Categorical(monthly_avg["month"],␣
 ↪categories=month_order, ordered=True)

sns.lineplot(data=monthly_avg, x="month", y="wind_speed", label="Wind Speed",␣
 ↪marker='o')
sns.lineplot(data=monthly_avg, x="month", y="gust", label="Gust", marker='o')
```

```
plt.title("Figure 2.1 Monthly Average Wind Speed & Gust (All States)")
plt.ylabel("Speed (m/s)")
plt.xlabel("Month")
plt.grid(True)
plt.tight_layout()
plt.show()
```



1.2.11 Insights:

Wind speed and gust intensity exhibit a strong positive correlation, indicating that periods of higher sustained winds also bring more intense gusts.

There is a notable dip in both metrics during March to April, suggesting a transitional weather period with lower wind activity.

From April to October, wind speeds and gusts show a steady climb, reflecting increasing atmospheric instability and possibly pre-monsoonal buildup.

A sharp decline is observed between October and November, likely marking the shift from the Southwest to the Northeast Monsoon.

This is followed by a rebound in December, aligning with the onset of the Northeast Monsoon.

1.2.12 Implications:

These patterns are strongly driven by monsoonal cycles, underlining the importance of regional climate systems in wind energy forecasting.

The seasonal dip in wind activity (March–April) may affect energy generation reliability and should be accounted for in resource planning.

The steady increase in wind speed from mid-year to October presents an optimal window for peak wind energy capture.

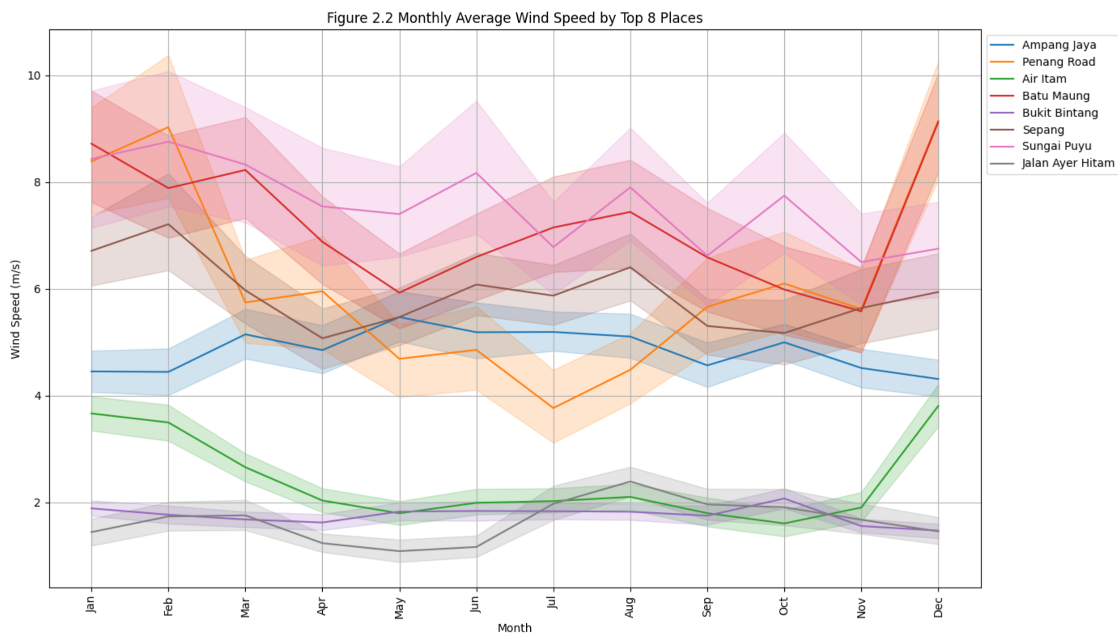
The post-monsoon rebound in December offers a second wind energy opportunity that can be strategically utilized.

Understanding these dynamics is essential for designing efficient wind energy systems and optimizing turbine deployment and maintenance schedules.

Monthly wind speed trend by place

[ ]:
```
month_order = ['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun',
               'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec']
df_clean_ws = df.dropna(subset=["wind_speed"])
df_clean_ws = df_clean_ws.copy()
df_clean_ws["month"] = pd.Categorical(df_clean_ws["month"],
 ↪categories=month_order, ordered=True)

# Step 1: Find places that have data for all 12 months
places_with_all_months = (
    df_clean_ws.groupby("place")["month"]
    .nunique()
    .loc[lambda x: x == 12]  # Keep only places with 12 unique months
    .index
)

# Step 2: Filter to places with full data
df_full_data = df_clean_ws[df_clean_ws["place"].isin(places_with_all_months)]

# Step 3: Get top 8 places by average wind speed among those
top_8_places = (
    df_full_data.groupby("place")["wind_speed"]
    .mean()
    .sort_values(ascending=False)
    .head(8)
    .index
)

# Step 4: Filter again for those top 8
df_top_8_full = df_full_data[df_full_data["place"].isin(top_8_places)]

# Step 5: Plot
```

```python
plt.figure(figsize=(14, 8))
sns.lineplot(data=df_top_8_full, x="month", y="wind_speed", hue="place",
    estimator="mean")
plt.legend(loc='upper left', bbox_to_anchor=(1, 1))
plt.title("Figure 2.2 Monthly Average Wind Speed by Top 8 Places")
plt.ylabel("Wind Speed (m/s)")
plt.xlabel("Month")
plt.xticks(rotation=90)
plt.grid(True)
plt.tight_layout()
plt.show()
```



1.2.13 Insights:

Penang exhibits greater monthly variability, consistent with it being a coastal region exposed to:

Sea breezes, Northeast and Southwest monsoon systems

Inland regions, such as Sepang, display more stable patterns throughout the year.

Wind behavior in Malaysia shows region-specific temporal dynamics, especially between coastal and inland areas.

1.2.14 Conclusion:

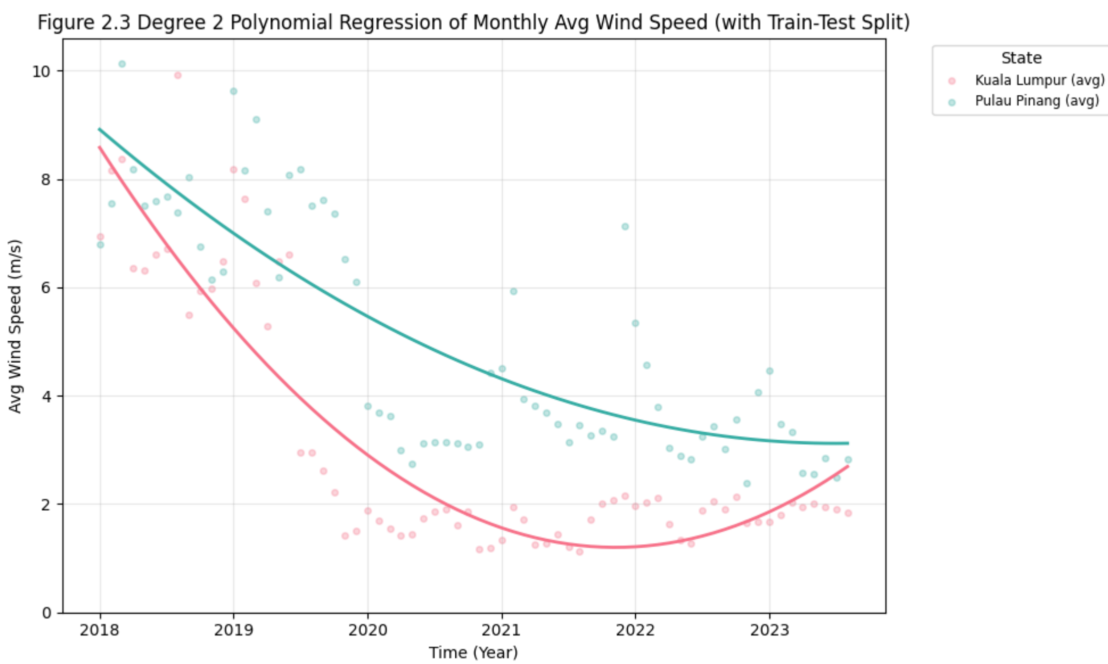
Sungai Puyu is the most reliable location in terms consistent high wind speed data.

Penang's high variability aligns with its exposure to coastal and seasonal weather patterns, making it a critical area for seasonal wind forecasting and planning.

[ ]:
```python
# Month mapping and numeric time
month_map = {
    'Jan': 1, 'Feb': 2, 'Mar': 3, 'Apr': 4, 'May': 5, 'Jun': 6,
    'Jul': 7, 'Aug': 8, 'Sep': 9, 'Oct': 10, 'Nov': 11, 'Dec': 12
}
df_states = df_clean_ws.copy()
df_states['month_num'] = df_states['month'].map(month_map).astype(int)

# Group to monthly averages
monthly_avg = (
    df_states.groupby(['state', 'year', 'month_num'])['wind_speed']
    .mean()
    .reset_index()
)
monthly_avg['time_numeric'] = monthly_avg['year'] + (monthly_avg['month_num'] -␣
↪1) / 12

# Set up plot
plt.figure(figsize=(10, 6))
states = monthly_avg['state'].unique()
palette = sns.color_palette("husl", n_colors=len(states))

print("Polynomial Regression (Degree 2) - Test Metrics by State (Monthly␣
↪Averages):\n")

for i, state in enumerate(states):
    data = monthly_avg[monthly_avg['state'] == state].copy()
    X = data['time_numeric'].values.reshape(-1, 1)
    y = data['wind_speed'].values

    # Train-test split (85% train, 15% test)
    X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.85,␣
↪random_state=42)

    # Fit polynomial regression
    poly = PolynomialFeatures(degree=2)
    X_train_poly = poly.fit_transform(X_train)
    X_test_poly = poly.transform(X_test)

    model = LinearRegression().fit(X_train_poly, y_train)
    y_test_pred = model.predict(X_test_poly)

    # Evaluate on test data
    mse = mean_squared_error(y_test, y_test_pred)
    r2 = r2_score(y_test, y_test_pred)
    print(f"{state}: Test MSE = {mse:.3f}, R² = {r2:.3f}")
```

```python
    # Plot original data and fit line (using full data)
    X_all_poly = poly.transform(X)
    y_all_pred = model.predict(X_all_poly)
    plt.scatter(X, y, label=f"{state} (avg)", color=palette[i], alpha=0.3, s=15)
    plt.plot(X, y_all_pred, color=palette[i], linewidth=2)

# Customize plot
plt.title("Figure 2.3 Degree 2 Polynomial Regression of Monthly Avg Wind Speed␣
↪(with Train-Test Split)")
plt.xlabel("Time (Year)")
plt.ylabel("Avg Wind Speed (m/s)")
plt.ylim(bottom=0)
plt.legend(title="State", bbox_to_anchor=(1.05, 1), loc='upper left',␣
↪fontsize='small')
plt.grid(True, alpha=0.3)
plt.tight_layout()
plt.show()
```

```
Polynomial Regression (Degree 2) - Test Metrics by State (Monthly Averages):

Kuala Lumpur: Test MSE = 1.565, R² = 0.761
Pulau Pinang: Test MSE = 1.577, R² = 0.710
```



1.2.15 Insights

U-shaped Trend Both States show a clear decline in average wind speed from 2018 through roughly mid-2021, followed by a modest rebound thereafter.

Kuala Lumpur falls more sharply from ~8.5 m/s down to ~1.2 m/s and then rises back toward ~2.5 m/s by 2023.

Pulau Pinang declines from ~9 m/s to ~3.2 m/s, then levels off around ~3.1 m/s.

State-to-State Differences Pulau Pinang consistently exhibits higher wind speeds (by ~1–2 m/s) than Kuala Lumpur, even though both follow the same general shape.

Variability (scatter) around the curves is larger in Pulau Pinang, suggesting more month-to-month fluctuation.

Timing of the Turning Point The minimum wind speeds occur around 2021–early 2022, indicating a multi-year low phase, likely tied to shifts in monsoon intensity or broader climatic anomalies.

A degree-2 polynomial captures these broad "fall then rise" patterns very effectively:

Kuala Lumpur: Test MSE = 1.565, $R^2 = 0.761$

Pulau Pinang: Test MSE = 1.577, $R^2 = 0.710$

1.2.16 Implications

Resource Planning The extended decline in wind speeds through mid-2021 would have eroded energy yields; developers should verify whether this was a one-off event or part of a longer climate trend.

The post-2021 uptick in Kuala Lumpur suggests potential recovery—an encouraging sign for new projects there—but Pulau Pinang's plateau warns against over-optimism.

Forecasting & Modeling A simple quadratic (degree-2) model captures the broad "fall then rise" pattern, but the fairly wide scatter—especially in Pulau Pinang—points to the need for seasonal or stochastic components in any production forecast.

Incorporating monsoon indices or El Niño/La Niña indicators could improve prediction of these multi-year swings.

Operational Strategy Maintenance and financing plans should account for multi-year lows; e.g., scheduling major overhauls outside the trough period to avoid compounding revenue drops.

Turbine siting and sizing might favor Pulau Pinang for baseline yield, but Kuala Lumpur could be a better bet if the rebound trend continues.

Risk Management The pronounced dip and rebound underscore the importance of long-term monitoring—relying on just 3–5 years of data risks missing these slow oscillations.

Diversification across multiple sites (and/or complementary energy sources) will buffer portfolios against region-specific lulls.

1.2.17 Research Question 3: "How does wind chill vary by region and time of day, and what does this imply about perceived temperature extremes?"

Qi Zhi

1.2.18 Solution:

[7]:
```
df = pd.read_csv('malaysia_weather_data.csv')

df_filtered_q3 = df.dropna(subset=['wind_chill', 'hour', 'place', 'state',␣
↪'wind_speed', 'temperature'])
df_filtered_q3.head()
```

[7]:
```
                    place          city         state  temperature  pressure  \
0  Bandar Sri Permaisuri  Kuala Lumpur  Kuala Lumpur         30.3  1015.325
1            Ampang Jaya  Kuala Lumpur  Kuala Lumpur         27.0  1000.680
2            Bukit Jalil  Kuala Lumpur  Kuala Lumpur         26.4  1013.550
3            Bukit Jalil  Kuala Lumpur  Kuala Lumpur         31.1  1005.420
4            Penang Road   George Town  Pulau Pinang         25.8  1007.790

   dew_point  humidity  wind_speed  gust  wind_chill  uv_index  \
0       25.1      73.7         0.9   1.5        30.3       1.0
1       25.6      92.0         5.2   5.3        27.0       0.0
2       26.3      99.0         0.0   0.0        26.4       NaN
3       30.9      99.0         0.0   0.0        31.1       NaN
4       23.9      88.0         6.4  16.1        25.8       NaN

   precipitation_rate  precipitation_total  year month  day  hour  minutes  \
0                 0.0                 0.00  2022   Sep   23    13       54
1                 0.0                17.78  2023   May    4    19        9
2                 0.0                32.00  2021   Oct   23    23       44
3                 0.0                 0.00  2021   Mar   21    18        9
4                 0.0                 0.00  2022   Dec   22     7        0

   seconds
0       57
1       59
2       43
3       47
4       23
```

1.2.19 Insights of wind chill distribution by region

[ ]:
```
# BOX PLOT FOR WIND CHILL DISTRIBUTION BY REGION
plt.figure(figsize=(12, 6))
df_filtered_q3.boxplot(column='wind_chill', by='place', rot=90)
plt.title('Figure 3.1 Wind Chill Distribution by Region')
plt.suptitle('')
```

```
plt.xlabel('Region')
plt.ylabel('Wind Chill (°C)')
plt.tight_layout()
plt.show()
```

```
<Figure size 1200x600 with 0 Axes>
```

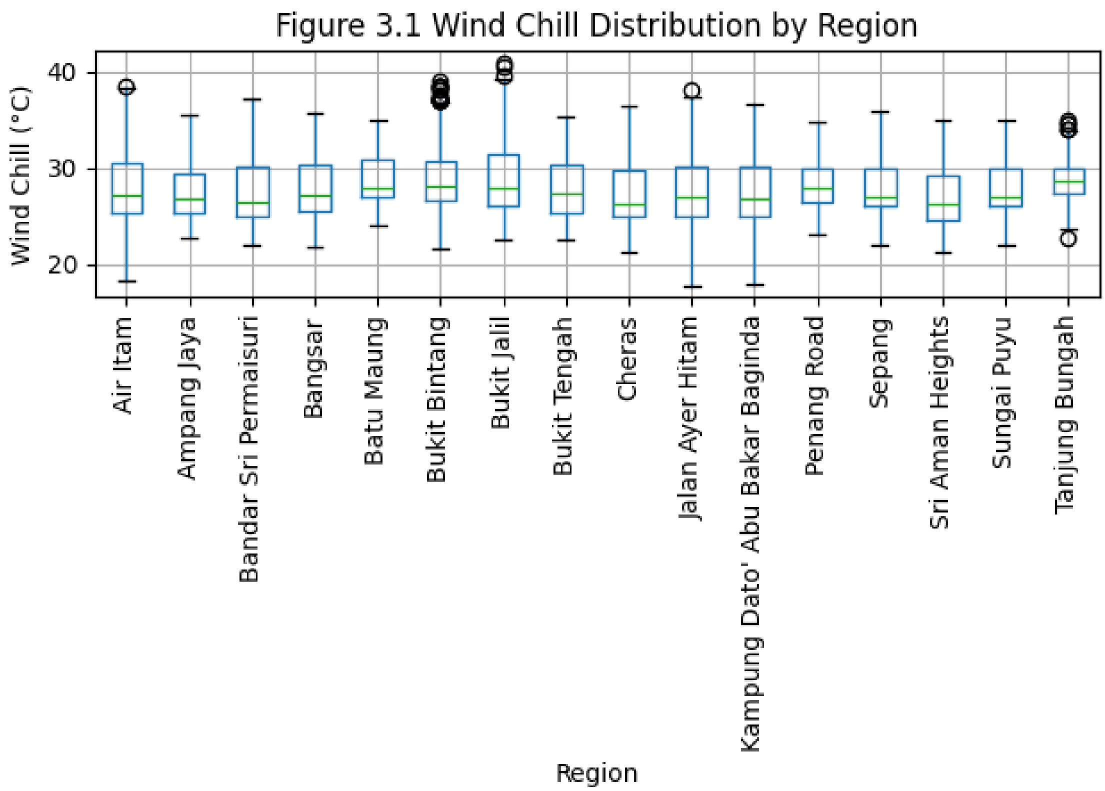


Figure 3.1 is a box plot that visualises data of the wind chill distribution throughout the regions. The following data is visualised for the concern that wind chill may impact turbine efficiency and thermal stress on the given systems.

1.2.20 Insights

- The suburbs have a wind chill with lower temperatures from the minimum values and show greater variability
- The big interquartile ranges indicate that there are great variability in wind chill across regions which implies that the wind chill are very out of the place due to elevation differences.

1.2.21 Implications

- Regions with lower medians and high variability are suggested to install thermal resilience in their wind turbine designs

- It is recommended to have wind installations in regions with lower variability and near to the coast where wind is higher.

1.2.22 Insights of average wind chill across hours of the day

[ ]:
```
# LINE PLOT FOR WIND CHILL ACROSS HOURS
plt.figure(figsize=(10, 5))
hourly_avg = df_filtered_q3.groupby('hour')['wind_chill'].mean()
plt.plot(hourly_avg.index, hourly_avg.values, marker='o')
plt.title('Figure 3.2 Average Wind Chill by Hour of Day')
plt.xlabel('Hour')
plt.ylabel('Average Wind Chill (°C)')
plt.grid(True)
plt.xticks(np.arange(0, 24, 1))
plt.tight_layout()
plt.show()
```

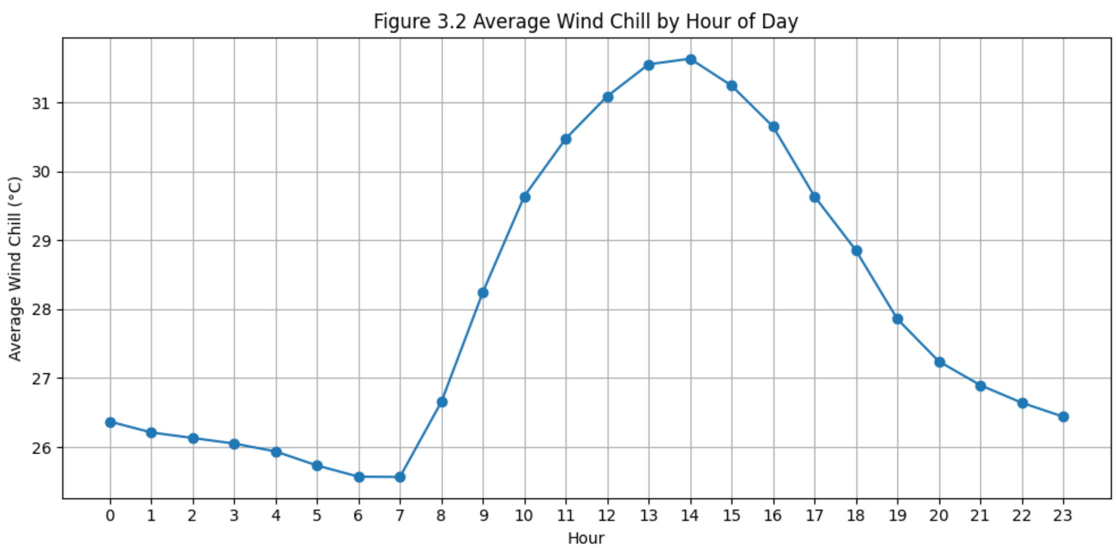


Figure 3.2 is a Line Plot of the average wind chill by the hour of the day. The plot may concern the diurnal variation affects energy use

1.2.23 Insights

- The line plot suggests that wind chill values are lowest between 3am to 7am having a minimum average of 25.5 degrees whilst peak values are seen at 1pm to 4pm having an average of 28 degrees.
- The range between the two averages is 3 degrees which are enough to affect sensitive equipment
- The standard deviation between hours is close to one degree celcius which tells us that there is a trend of variation throughout a day which can be predicted.

1.2.24 Implications

- Early morning hours are colder as given in the insights which implies that maintenance operations need to be taken in to consideration where low temperatures may increase risk of condensation on electrical equipments.
- With the standard deviation being low it allows for predictive scheduling in maintenance

1.2.25 Insights of wind speed against wind chill

[ ]:
```python
# SCATTER PLOT FOR WIND SPEED AGAINST WIND CHILL
plt.figure(figsize=(10, 6))
plt.scatter(df_filtered_q3['wind_speed'], df_filtered_q3['wind_chill'], alpha=0.
↪5)
plt.title('Figure 3.3 Wind Speed vs Wind Chill')
plt.xlabel('Wind Speed (km/h or m/s)')
plt.ylabel('Wind Chill (°C)')
plt.grid(True)
plt.tight_layout()
plt.show()
```

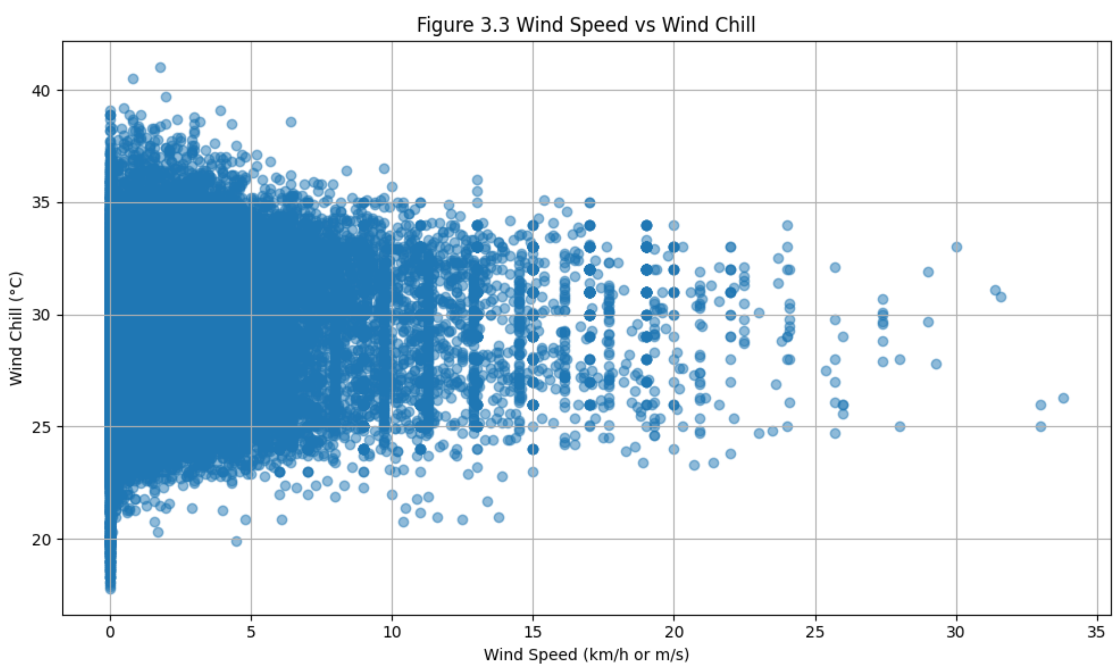


Figure 3.3 shows a scatter plot where wind speed is plotted against wind chill. This plot may give insights as to whether there is a need to factor in thermal design for high wind areas.

1.2.26 Insights

- The scatter plot shows a slightly negative trend where as wind speed increases,wind chill tends to decrease but note that the correlation is weak.

1.2.27 Implications

In regions where there is a lot of wind, temperature may not be excessive, but prolonged exposure to cold winds may affect sensitivity of equipment

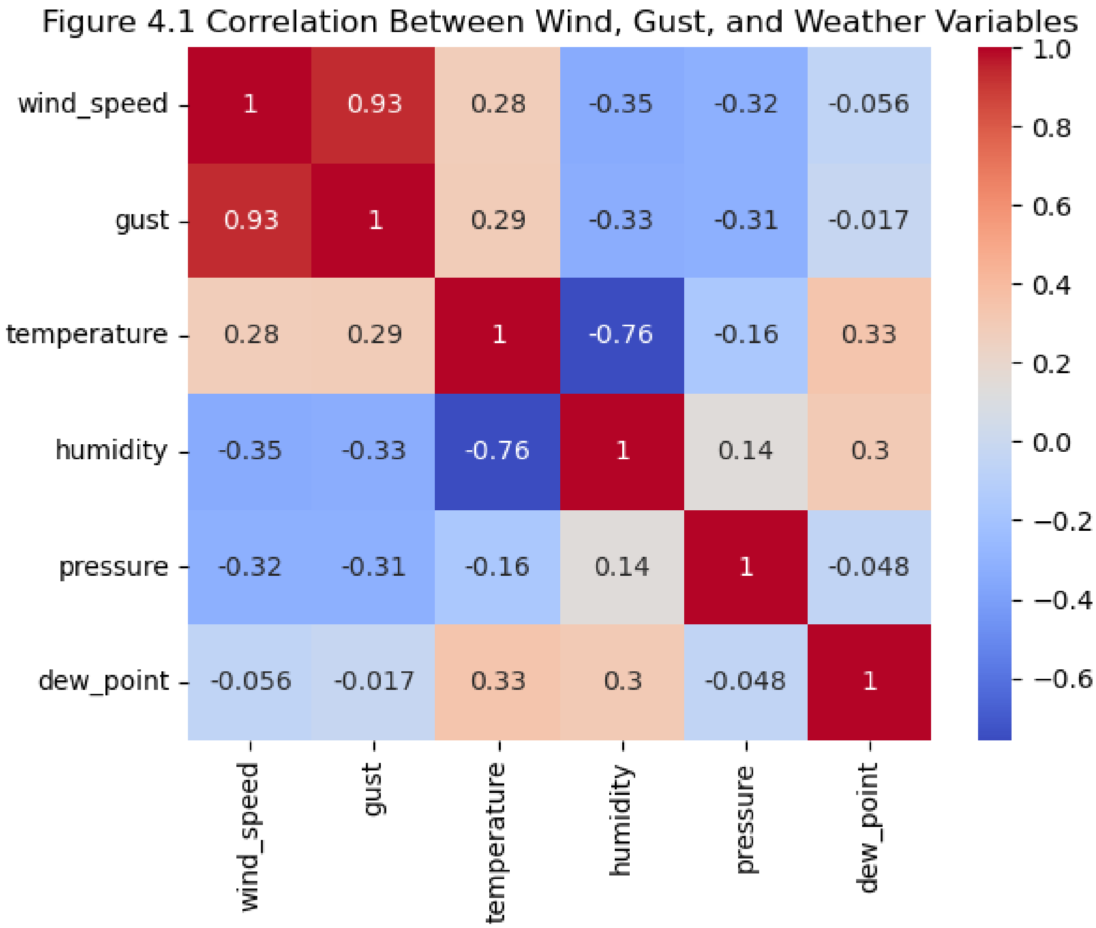
1.2.28 Research Question 4: "What is the relationship between wind speed, gust, and other weather variables like humidity, temperature, and pressure?"

Sharon

1.2.29 Solution:

1.2.30 Heatmap gives a quick overview of how strongly each pair of variables is correlated

[8]:
```
corr = df[['wind_speed', 'gust', 'temperature', 'humidity', 'pressure',␣
 ↪'dew_point']].corr()
sns.heatmap(corr, annot=True, cmap='coolwarm')
plt.title('Figure 4.1 Correlation Between Wind, Gust, and Weather Variables')
plt.show()
```



The code above generates a correlation heatmap to visualise the strength and direction of relationship between key weather variables: wind speed, gust, temperature, humidity, pressure, and dew point. The correlation matrix is calculated using Pearson's correlation coefficient, which ranges from -1 (perfect negative correlation) to +1 (perfect positive correlation), with 0 indicating no linear relationship.

From the heatmap, we can observe that: 1. Wind speed and gust show a strong positive correlation (0.93), which is expected as gust is often an extension of wind speed peaks. 2. Both pressure and humidity show weak or negative correlations with wind speed and gust, indicating that these variables may not strongly influence wind behavior in the observed data. 3. Temperature shows a weak positive correlation with both wind speed and gust, suggesting that as temperature slightly increases, wind speed and gust may also tend to increase — but the relationship is not strong. 4. The dew point shows a very weak negative correlation with both wind speed and gust. These near-zero values indicate that there is virtually no linear relationship between dew point and wind behavior. In other words, changes in dew point do not meaningfully correspond to changes in wind speed or gust intensity in the data analyzed.

1.2.31 Linear Regression Scatter Plot (Wind Speed vs Gust)

[ ]:
```
sns.lmplot(data=df, x='gust', y='wind_speed', scatter_kws={'alpha': 0.3,␣
↪'color': 'darkorange'}, line_kws={'color': 'navy'})
plt.title('Figure 4.2 Wind Speed vs Gust')
```

[ ]:
```
Text(0.5, 1.0, 'Figure 4.2 Wind Speed vs Gust')
```

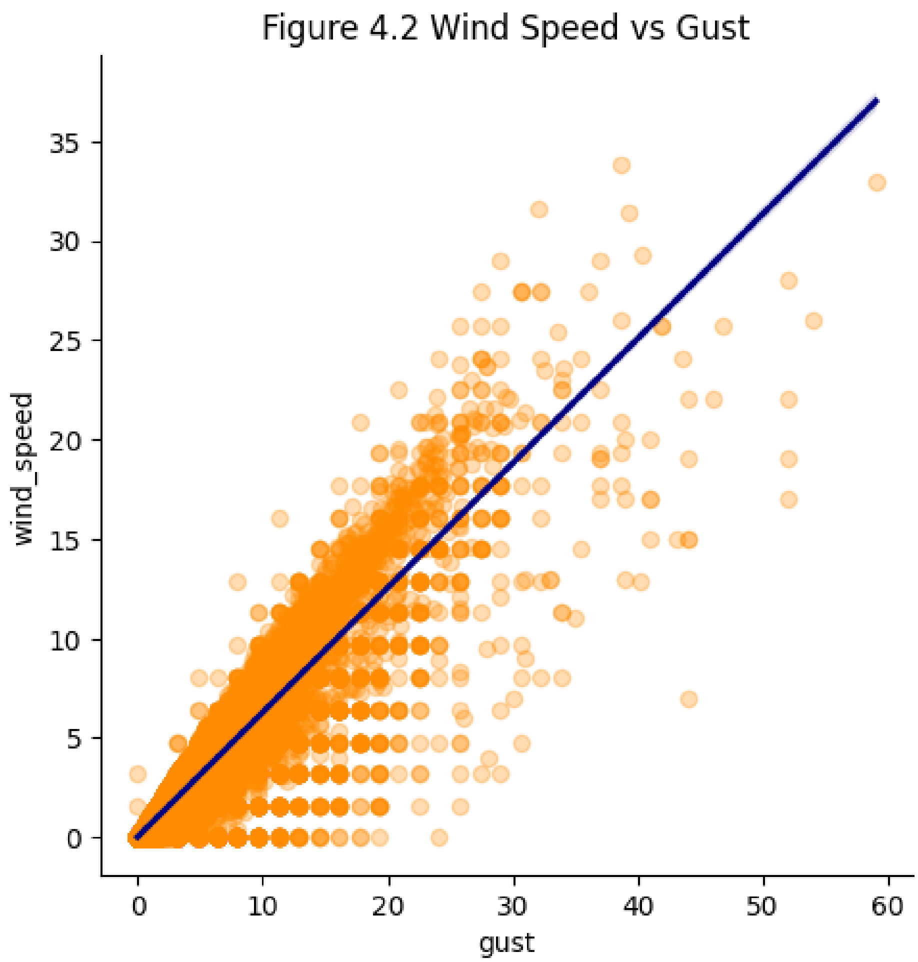


Figure 4.2 shows a clear positive linear relationship between gust and wind speed. As gust values increase, wind speed generally increases as well. The data points form a dense cluster along the diagonal regression line, especially between gust values of 0 to 30, indicating that wind speed tends to rise proportionally with gust in this range. Although there is more spread at higher gust values (above 30), the overall trend remains consistent.

While the majority of data points in the scatter plot follow a strong linear trend between gust and wind speed, a few outliers are visible. These are points where gust values are unusually high but do not correspond with a similar increase in wind speed. These outliers may be due to anomalies in weather conditions, such as isolated gust events, or could reflect measurement inaccuracies.

Acknowledging these outliers is important, as they slightly reduce the precision of the regression model and highlight the need to consider data quality and environmental context in real-world datasets.

1.2.32 Linear Regression Scatter Plot (Wind Speed vs Temperature)

[ ]:
```
sns.lmplot(data=df, x='temperature', y='wind_speed',scatter_kws={'alpha': 0.3,␣
↪'color': 'gray'}, line_kws={'color': 'blue'})
plt.title('Figure 4.3 Wind Speed vs Temperature')
plt.xlabel('Temperature (°C)')
plt.ylabel('Wind Speed')
plt.show()
```

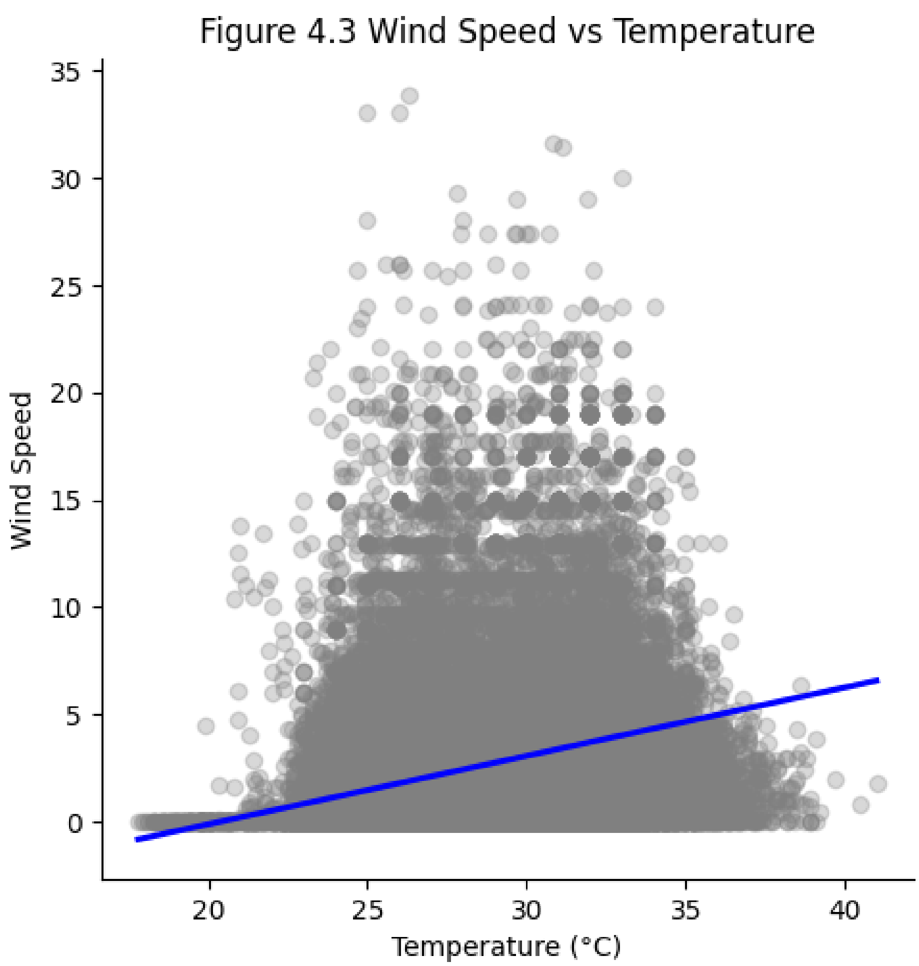


Figure 4.3 shows a weak positive relationship between temperature and wind speed, consistent with the correlation value of around 0.28, indicating only a slight tendency for wind speed to increase as temperature rises. The data points are highly concentrated between 20°C to 35°C, which reflects Malaysia's tropical climate. However, the spread of wind speed within this temperature range is

quite wide, suggesting that temperature alone is not a strong predictor of wind speed.

There are also a few outliers where wind speed exceeds 30 even at moderate temperatures, which may result from unusual weather events or data logging inconsistencies. These outliers should be noted, as they can slightly skew the regression line.

1.2.33 Linear Regression Scatter Plot (Wind Speed vs Humidity):

We can observe a weak negative correlation, which means that as humidity increases, wind speed tends to decrease slightly. It can be used to contrast with the stronger relationship.

[ ]:
```
sns.lmplot(data=df, x='humidity', y='wind_speed', scatter_kws={'alpha': 0.3,␣
↪'color': 'gray'}, line_kws={'color': 'navy'})
plt.title('Figure 4.4 Wind Speed vs Humidity')
```

[ ]:
```
Text(0.5, 1.0, 'Figure 4.4 Wind Speed vs Humidity')
```

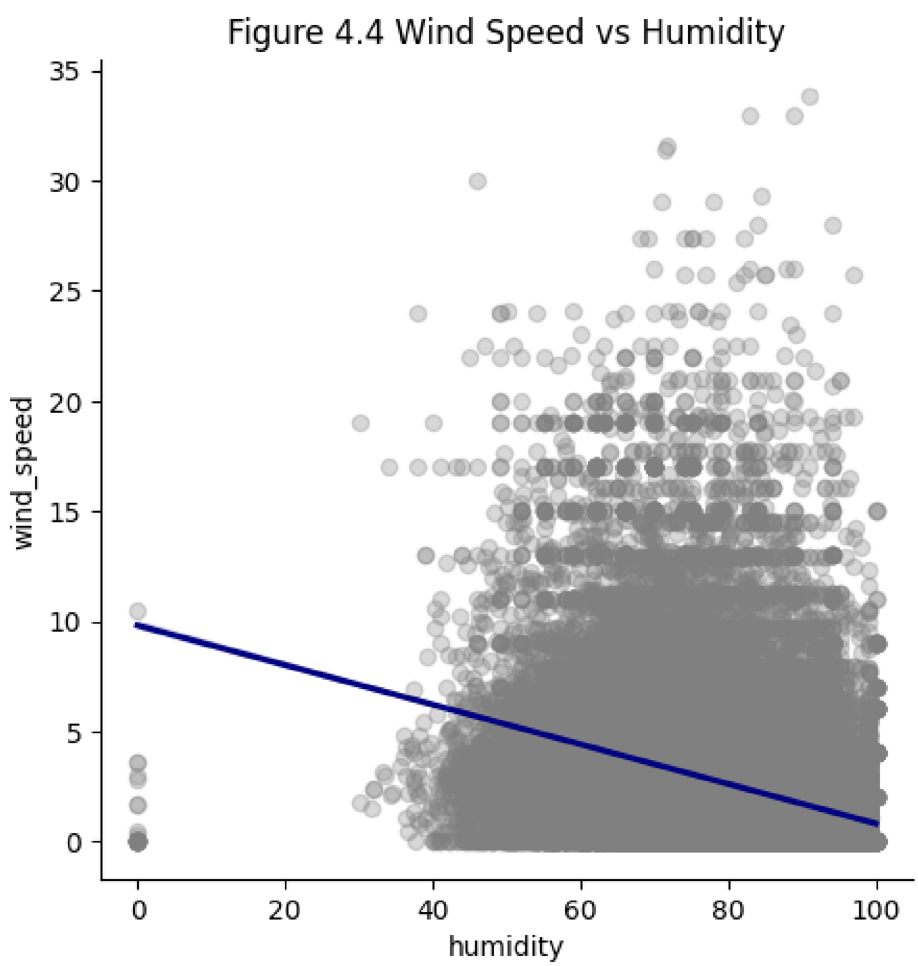


Figure 4.4 shows a weak negative correlation between humidity and wind speed, which aligns with

the slightly downward trend in the regression line. As humidity increases, wind speed tends to decrease slightly, although the relationship is not strong. The data is densely clustered at higher humidity levels, with wind speed mostly below 20 km/h. A few high wind speed outliers could be due to localized weather conditions or errors in data collection.

1.2.34 Conclusion

The graphs show that wind speed and gust have a strong positive correlation, making gust a useful indicator for wind energy potential. Temperature has a weak positive relationship with wind speed, while humidity shows a weak negative correlation. These findings suggest that areas with frequent gusts, moderate temperatures, and lower humidity may be more suitable for wind energy development in Malaysia.

1.2.35 Research Question 5: "Which regions exhibit the most extreme wind events (high gusts or wind speed) and how often do these occur?"

Anusha

1.2.36 Solution:

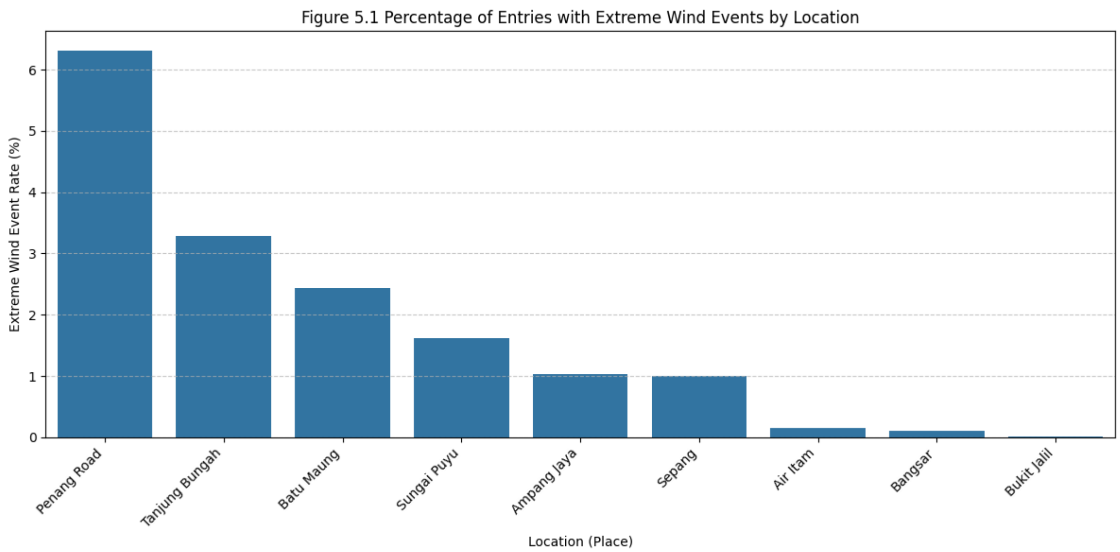
[ ]:
```
# REMOVE VALUES WITH BOTH WIND SPEED AND GUST MISSING
df_clean5 = df[~(df['wind_speed'].isna() & df['gust'].isna())]
df_clean5.head()
```

[ ]:
```
                  place          city         state  temperature  pressure  \
0  Bandar Sri Permaisuri  Kuala Lumpur  Kuala Lumpur         30.3  1015.325
1            Ampang Jaya  Kuala Lumpur  Kuala Lumpur         27.0  1000.680
2            Bukit Jalil  Kuala Lumpur  Kuala Lumpur         26.4  1013.550
3            Bukit Jalil  Kuala Lumpur  Kuala Lumpur         31.1  1005.420
4            Penang Road   George Town  Pulau Pinang         25.8  1007.790

   dew_point  humidity  wind_speed  gust  wind_chill  uv_index  \
0       25.1      73.7         0.9   1.5        30.3       1.0
1       25.6      92.0         5.2   5.3        27.0       0.0
2       26.3      99.0         0.0   0.0        26.4       NaN
3       30.9      99.0         0.0   0.0        31.1       NaN
4       23.9      88.0         6.4  16.1        25.8       NaN

   precipitation_rate  precipitation_total  year month  day  hour  minutes  \
0                 0.0                 0.00  2022   Sep   23    13       54
1                 0.0                17.78  2023   May    4    19        9
2                 0.0                32.00  2021   Oct   23    23       44
3                 0.0                 0.00  2021   Mar   21    18        9
4                 0.0                 0.00  2022   Dec   22     7        0

   seconds
0       57
1       59
```

```
2         43
3         47
4         23
```

1.2.37 Histogram for High Wind-Speed and Gust

```
[ ]: HIGH_WIND_SPEED = 20 # SET BOUNDARY TO DECLARE EXTREME WIND-SPEED
     HIGH_GUST = 25 # SET BOUNDARY TO DECLARE EXTREME GUST

     # FILTER ROWS WITH EXTREME WIND-SPEED AND GUST
     extreme_wind_df = df_clean5[(df_clean5['wind_speed'] >= HIGH_WIND_SPEED) |␣
      ↪(df_clean5['gust'] >= HIGH_GUST)]

     # REPLACE MISSING place VALUES WITH "Unknown"
     extreme_wind_df = extreme_wind_df.dropna(subset=['place'])

     # TOTAL ENTRIES PER LOCATION
     total_days_per_place = df_clean5['place'].value_counts()

     # TOTAL EXTREME WIND EVENTS PER LOCATION
     extreme_days_per_place = extreme_wind_df['place'].value_counts()

     # CALCULATE THE PERCENTAGE OF DAYS WITH EXTREME WIND CONDITIONS
     extreme_event_percentage = (extreme_days_per_place / total_days_per_place) * 100
     extreme_event_percentage = extreme_event_percentage.dropna().
      ↪sort_values(ascending=False)

     # PLOTTING THE FREQUENCY HISTOGRAM
     plt.figure(figsize=(12, 6))
     sns.barplot(x=extreme_event_percentage.index, y=extreme_event_percentage.values)
     plt.xticks(rotation=45, ha='right')
     plt.title("Figure 5.1 Percentage of Entries with Extreme Wind Events by␣
      ↪Location")
     plt.xlabel("Location (Place)")
     plt.ylabel("Extreme Wind Event Rate (%)")
     plt.tight_layout()
     plt.grid(axis='y', linestyle='--', alpha=0.7)
     plt.show()
```



PURPOSE:

Figure 5.1 illustrates the percentage of entries with extreme wind events. - The purpose is to pinpoint which regions exhibit the extreme natural wind activity, *most frequently*. - Helps identify priority zones for carefully targeted wind energy infrastructure development.

RELEVANCE:

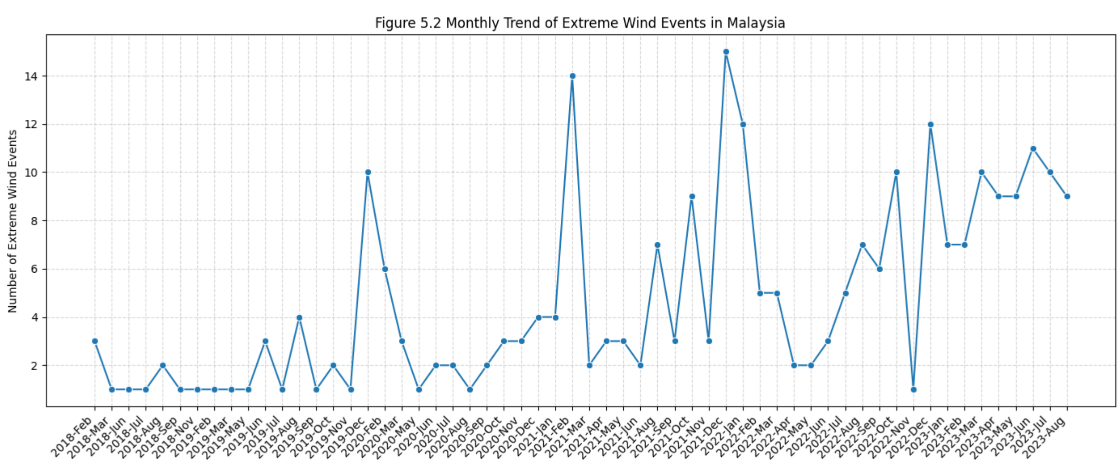
Figure 5.1 contributes directly to the core research question by revealing the regional likelihood of extreme wind events. - These spatial insights are critical for assessing where wind energy infrastructure can be most effective. - Helps guide Malaysia's renewable energy strategy through a geographically-informed lens.

KEY INSIGHTS: - Uneven distribution of wind events as seen in Figure 5.1 indicate that not all regions in Malaysia are equally viable for wind energy. - Top three locations with the highest number of extreme wind events are: 1. Penang Road 2. Tanjung Bungah 3. Batu Maung - These regions show consistently high wind activity, hence targeted planning towards these 3 regions.

STAKEHOLDER IMPLICATIONS: - *Government Agencies*: - Can map wind risk zones to identify where safe deployment is feasible. - Invest in resilient infrastructure policies and fund R&D for turbines adapted to Malaysia's wind climate. - Use meteorological data to support real-time grid-balancing and forecasting systems. - *Renewable Energy Companies*: - Opportunity to test or showcase high-durability turbines adapted for Southeast Asia. - Use wind frequency data to develop revenue models tied to seasonal wind patterns. - Focus on insurance-backed investment in risk-managed zones. - *Urban Planners & Housing Development Authorities*: - Design energy zoning frameworks that integrate wind infrastructure into non-residential land (e.g., coastlines, ridgelines, industrial parks). - Align planning permissions with extreme weather safety standards. - *Research Institutions & Tech Innovators*: - Innovate turbine materials, aerodynamics, and predictive AI models that adapt to Malaysia's wind extremes. - Pilot smart grid systems in these regions that adjust power flows dynamically.

1.2.38 Line Plot of Extreme Wind Events Over Time

[ ]:
```python
# GROUP BY YEAR AND MONTH, AND COUNT THE NUMBER OF EXTREME WIND EVENTS PER GROUP
extreme_wind_and_gust_df = df_clean5[(df_clean5['wind_speed'] >= HIGH_WIND_SPEED) | (df_clean5['gust'] >= HIGH_GUST)]
monthly_counts = (extreme_wind_and_gust_df.groupby(['year', 'month']).size().reset_index(name='extreme_event_count'))

# CONVERT ABBREVIATED MONTH NAMES TO NUMERICAL FORMAT (Jan = 1, Feb = 2...)
monthly_counts['month_num'] = pd.to_datetime(monthly_counts['month'], format='%b', errors='coerce').dt.month

# REMOVE ROWS WITH INVALID/MISSING MONTH VALUES
monthly_counts = monthly_counts.dropna(subset=['month_num'])

# SORT DATA CHRONOLOGICALLY BY YEAR AND NUMERICAL MONTH
monthly_counts = monthly_counts.sort_values(['year', 'month_num'])

# CREATE A NEW COLUMN TO USE ON THE X-AXIS IN THE "YYYY-MMM" FORMAT
monthly_counts['year_month'] = monthly_counts['year'].astype(str) + '-' + monthly_counts['month']

# PLOT THE LINE GRAPH
plt.figure(figsize=(14, 6))
sns.lineplot(data=monthly_counts, x='year_month', y='extreme_event_count', marker='o')
plt.title("Figure 5.2 Monthly Trend of Extreme Wind Events in Malaysia")
plt.xlabel("Year-Month")
plt.ylabel("Number of Extreme Wind Events")
plt.xticks(rotation=45, ha='right')
plt.grid(True, linestyle='--', alpha=0.5)
plt.tight_layout()
plt.show()
```



PURPOSE:

Figure 5.2 tracks the number of extreme wind events over time (from January 2018 to August 2023), based on monthly counts. - The purpose is to identify temporal trends, spikes, and patterns in wind behavior. - Helps identify predictability and planning of wind energy infrastructure.

RELEVANCE:

Figure 5.2 contributes directly to the core research question by highlighting how extreme wind events vary over time. - These insights reinforce the viability and urgency of integrating wind energy into Malaysia's renewable energy portfolio. - Understanding these temporal trends enables better policy and investment decisions in the energy sector.

KEY INSIGHTS: - Figure 5.2 demonstrates frequent fluctuations in wind conditions throughout the recorded period, demonstrating a rising trend. - Significant spikes are observed during: - November 2019 to February 2020 - January to March 2021 - December 2021 to January 2022 - December 2022 to January 2023 - These could indicate specific stormy months or extreme weather episodes during those periods.

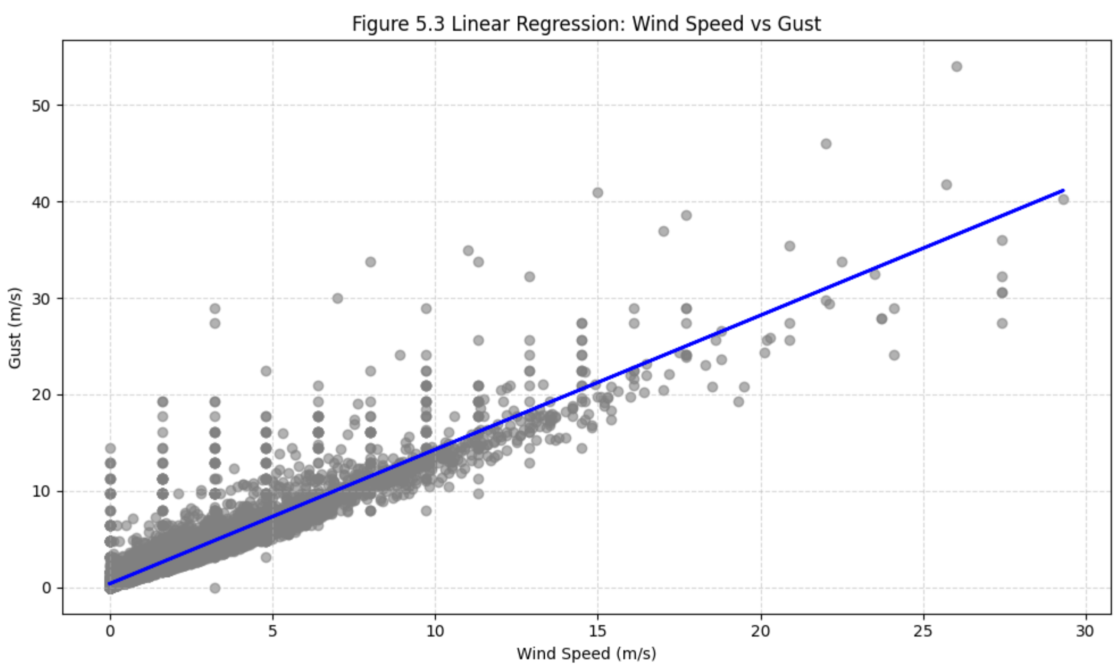
STAKEHOLDER IMPLICATIONS: - *Government Agencies*: - Need to reassess Malaysia's renewable energy targets and adapt grid infrastructure for sudden wind activity. - May prompt updates to climate risk forecasts, influencing long-term energy and environmental policy. - *Renewable Energy Companies*: - Fluctuating but increasing events offer opportunities for energy generation. - Also calls for flexible wind-turbine scheduling and energy-storage strategies. - Highlights the importance of seasonal wind forecasting models in operational planning. - *Policy-Makers*: - Seasonal surges suggest policy incentives could be timed with high-wind periods to maximize efficiency: - Tax breaks for renewable energy companies/private sector investors. - Issue tradable credits (e.g. Renewable Energy Certificates (RECs)) for energy generated during peak wind periods. - This allows renewable wind-energy companies to sell these credits to companies seeking to demonstrate that they meet sustainability targets and adhere to their CSR whilst producing their goods.

1.2.39 Linear Regression Between Wind Speed and Gust

[ ]:
```
# DROP ALL MISSING VALUES IN wind_speed AND gust COLUMNS
df_clean_linreg = df.dropna(subset=["wind_speed", "gust"])

# SELECT INDEPENDENT (wind_speed) AND DEPENDENT (gust) VARIABLE
X = df_clean_linreg[['wind_speed']]
y = df_clean_linreg['gust']

# SPLIT DATA INTO TRAINING AND TESTING SETS
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
↪random_state=42)

# INITIALIZE AND TRAIN LINEAR REGRESSION MODEL
model = LinearRegression()
model.fit(X_train, y_train)
```

```python
# PREDICT GUST VALUES FOR THE TEST SET
y_pred = model.predict(X_test)

# PRINT NOTABLE METRICS
print("MSE:", mean_squared_error(y_test, y_pred))
print("R²:", r2_score(y_test, y_pred))

# PLOT THE ACTUAL VS PREDICTED GUST VALUES
plt.figure(figsize=(10, 6))
plt.scatter(X_test, y_test, color='gray', alpha=0.6)
plt.plot(X_test, y_pred, color='blue', linewidth=2)
plt.xlabel("Wind Speed (m/s)")
plt.ylabel("Gust (m/s)")
plt.title("Figure 5.3 Linear Regression: Wind Speed vs Gust")
plt.grid(True, linestyle='--', alpha=0.5)
plt.tight_layout()
plt.show()
```

```
MSE: 2.747046609324536
R²: 0.8689799838031326
```



PURPOSE:

Figure 5.3 illustrates the relationship between wind speed and gust intensity in Malaysia using a linear regression model, and assesses how strongly wind speed predicts gust strength.

RELEVANCE:

- This result validates that gusts are closely tied to sustained wind speeds, meaning that regions with consistent wind can be reliably modeled for gust risk as well.
- This has important implications for wind energy planning, as turbines must account for both variables to prevent damage and ensure efficiency.

KEY INSIGHTS: - A strong positive linear relationship is visible: as wind speed increases, gust intensity generally increases as well. - The R^2 value of ~0.869 indicates that ~87% of the variation in gust values is explained by wind speed, suggesting a highly predictive relationship. - Some outliers appear with very high gusts at lower wind speeds, likely due to short-duration spikes or sensor anomalies. - The MSE value of ~2.75 shows relatively low error between predicted and actual values, confirming the model's strength.

STAKEHOLDER IMPLICATIONS: - *Government Agencies*: - May adopt wind speed as a proxy metric for national gust hazard forecasting, allowing simpler regulation and zoning. - *Renewable Energy Companies*: - Can use wind speed data to model and anticipate gust behavior, improving turbine design and siting strategies. - *Urban Planners*: - Areas with correlated wind and gust patterns may need reinforced zoning laws around infrastructure and turbine deployment. - *Turbine Manufacturers*: - Can integrate gust-tolerant designs in regions with steeper gust-to-wind-speed gradients.

1.2.40 Conclusion

Regions like Penang Road, Tanjung Bungah, and Batu Maung exhibit the highest frequency of extreme wind events, highlighting them as prime zones for wind energy development. The histogram reveals clear regional disparities, while the line plot shows variations in our trend, highlighting extreme wind events from 2018 to 2023, with evident spikes in extreme wind events during specific seasonal windows. The linear regression confirms a strong correlation between wind speed and gust intensity, allowing predictive modeling. Together, these findings emphasize both where and when extreme wind conditions occur, offering strategic insight for siting, policy planning, and renewable infrastructure tailored to Malaysia's wind landscape.

1.2.41 Research Question 6: "At what times of day do peak wind speeds and gusts typically occur across regions?"

Jia Yi

1.2.42 Solution:

[ ]:
```
#data cleaning
df_filtered_Q6 = df.dropna(subset=["gust", "hour", "state", "state",␣
 ↪"wind_speed"])
```

1.2.43 Line plot of average wind speed at KL and Pulau Pinang by hour

[ ]:
```
#create a pivot table for line graph data
pivot_wind = df_filtered_Q6.pivot_table(index='hour', columns='state',␣
 ↪values='wind_speed', aggfunc='mean').reset_index()
```

```
melted_pivot_wind = pivot_wind.melt(id_vars= 'hour', var_name='state',␣
    ↪value_name='wind_speed')
```

[ ]:
```
#line graph plotting
sns.relplot(data=melted_pivot_wind, x='hour', y='wind_speed', hue='state',␣
    ↪kind='line')
plt.title("Figure 6.1: Average wind speed at each hour of the day")
plt.xlabel('Each hour of a day')
plt.ylabel('Average wind speed (m/s)')
plt.xticks(range(0,24,2))
plt.xlim(0,23)
plt.grid(True)
plt.plot()
```

[ ]: []

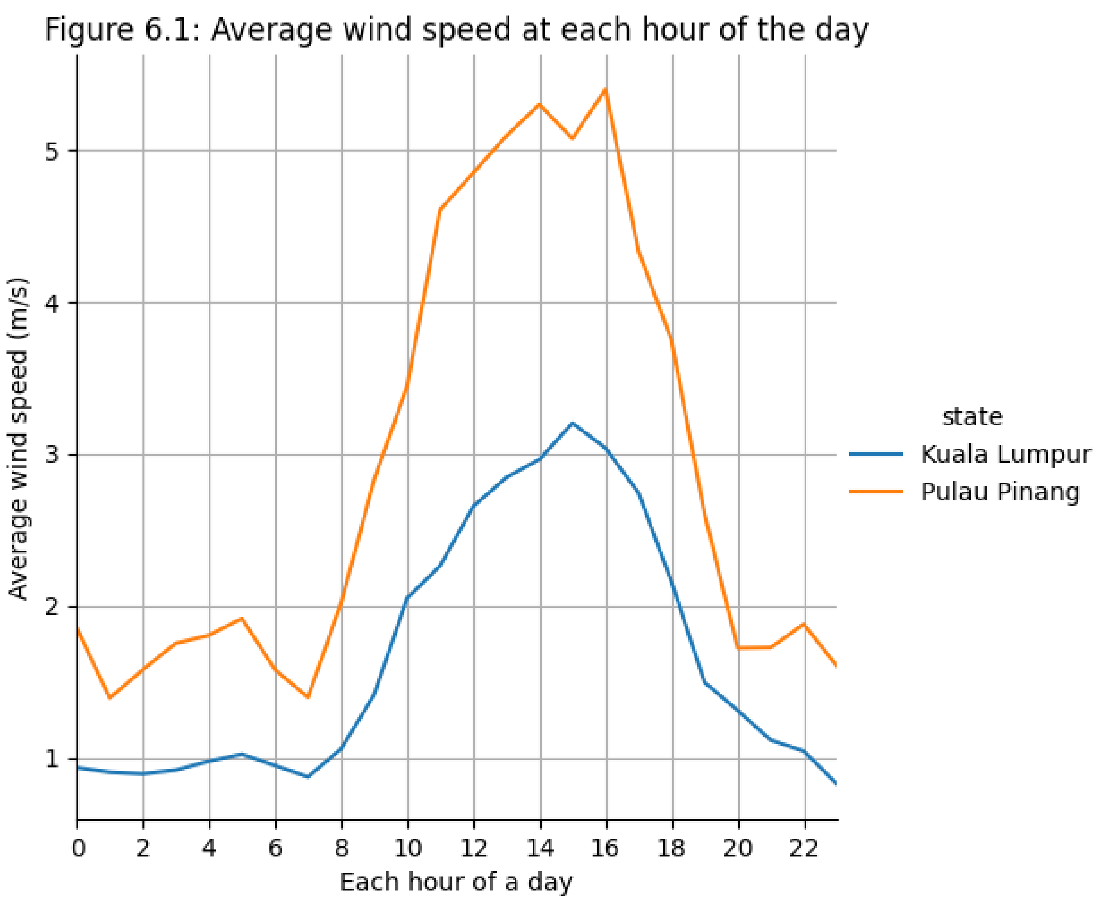


Figure 6.1 displays the average wind speed accross different hour of the day for each states. The x-axis represents the hour, and the y-axis represents the average wind speed (m/s). The orange line represents Pulau Pinang; blue line represents Kuala Lumpur.

From Figure 6.1, we can see that the peak occur at around 12:00PM and 4:00PM in Pulau Pinang, with speeds approaching 5.4m/s. This mid-afternoon rise aligns with increased solar heating and coastal breeze effects. While in KL, the peak occurs at around 3:00PM at just over 3.0m/s. There's a rise between 10:00AM-4:00PM, but still shows lower wind speeds compare to Pulau Pinang. However, both states reflect a consistent daily wind pattern which the wind speeds are lowest during the midnight (12AM-6AM), rise sharply at early morning (7AM) and taper off by evening (after 5-6PM).

1.2.44 Heat Map of wind speed at each hour throughout the day

[ ]:
```
#create a pivot table for heatmap data
q6_2_pivot = df_filtered_Q6.pivot_table(index='state', columns='hour',␣
↪values='wind_speed', aggfunc='mean')
```

[ ]:
```
# heatmap plotting
plt.figure(figsize=(q6_2_pivot.shape[1], q6_2_pivot.shape[0]))
sns.heatmap(q6_2_pivot, annot=True, cmap='coolwarm')
plt.xlabel("Hour of Day")
plt.ylabel("Region")
plt.title("Figure 6.2: Wind speed of each hour of the day")
plt.show()
```

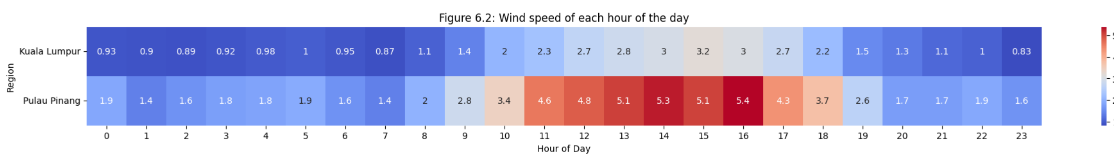


Figure 6.2 displays the average wind speed accross different hour of the day for each states. The x-axis represents the hour, and the y-axis represents the states; each boxes indicate the average wind speed at that particular time and corresponding state.

From Figure 6.2, we can observe that both Pulau Pinang and KL experience peak wind speed around the same time of the day (11:00AM-6:00PM), but Pulau Pinang shows stronger wind intensity while KL has weaker but cosistent wind intensity during that time. This visual highlighted inter-regional variability. Coastal region like Pulau Pinang shows stronger and more sustain wind speed; while urban centres like Kuala Lumpur's wind are weaker, likely due to urban obstacle (buildings distrupt the wind flow).

1.2.45 Stacked Bar Graph of Hourly Gust Occurrence at each State

[ ]:
```
#create pivot for bar chart
gust_count = df_filtered_Q6[['state','hour']].groupby(['state', 'hour']).size().
↪reset_index(name='count')
state_total = gust_count.groupby('state')['count'].sum().
↪reset_index(name='total_count')
gust_count = gust_count.merge(state_total, on='state')
```

```
gust_count['percent'] = (gust_count['count']/gust_count['total_count']) *100
q6_3_pivot= gust_count.pivot(index='hour', columns='state', values='percent')
```

[ ]:

```
#bar chart plotting
q6_3_pivot.plot(kind='bar', figsize=(10,4))
plt.title("Figure 6.3: Hourly Gust Occurrence by Regions (in Percentage)")
plt.xlabel("Hour of Day")
plt.ylabel("Percentage of Gusts (%)")
plt.xticks(rotation=0)
plt.legend(title="State", bbox_to_anchor=(1.05, 1), loc='upper left')
plt.tight_layout()
plt.show()
```

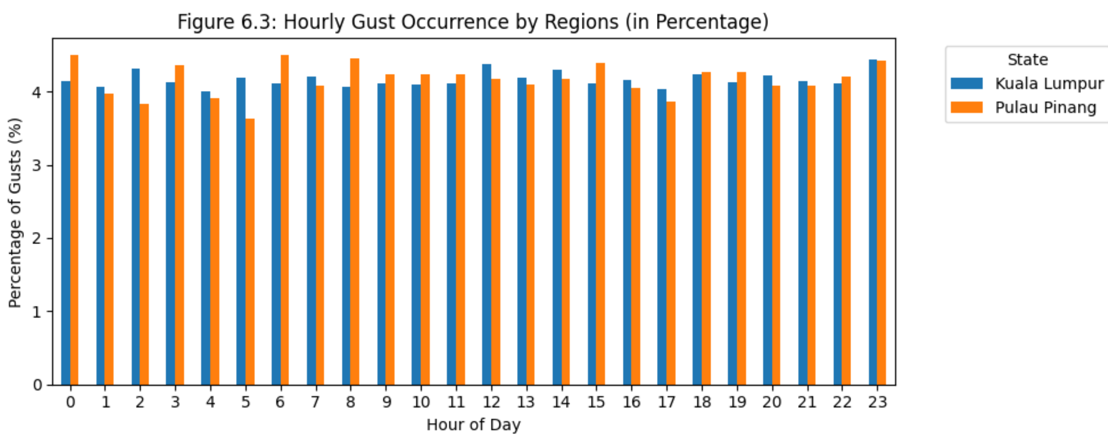


Figure 6.3 visualizes the total number of wind gust occurrences across each hour of the day, categorized by region. The x-axis represents the hour, y-axis represents the frequency of gust events. The orange bar represents Pulau Pinang, and the blue bar represents Kuala Lumpur.

This graph shows that Pulau Pinang has higher variability (higher highs and lower lows throughout the day) and KL is more stable, likely due to Penang being coastal and influenced by sea breezes, open areas, land-sea temp difference etc. Pulau Pinang's gust peaks are distribute around midnight (0–1AM), early morning (6–8AM), afternoon (3PM) & late night (11PM) — likely influenced by land-sea breeze interactions or dynamic coastal wind patterns. Kuala Lumpur shows more consistent gust distribution, with slight increases in early night (7–8PM) and late night (1AM). This lower variation likely due to urban obstacles in the city area.

There's no dominant gusting hour, suggesting that Malaysia's wind gusts are more evenly distributed than driven by sharp diurnal cycles.

1.2.46 Implications for Renewable Energy Sector

i) Wind Energy Planning & Turbine Scheduling

insights from graphs: Pulau Pinang shows stronger, more consistent afternoon wind speeds; Kuala Lumpur shows higher gust frequency, with milder average speeds.

Examples:

TNB Renewables and Petronas Clean Energy can priotize Pulau Pinang for pilot onshore or coastal wind farm development.

ii) Region-Specific Infrastructure Development

insights from graph: Different region require different wind energy solutions; uniform infrastructure rollout is inefficient

Examples:

Kuala Lumpur City Hall (DBKL) can explore small-scale urban wind turbines for smart buildings and rooftop use. Penang State Government and MBPP(Penang Island City Council) can allocate coastal land for mid-scale wind turbine installations. Malaysia Green Building Council(MGBC) can integrate region-specific wind data into sustainable building codes.

iii) Forecasting & Smart Grid Integration

insights from graph: Wind patterns are predictable by hour and region; data supports real-tume and long-term forecasting

Examples:

TNB Grid Services can integrate hourly wind predictions into smart grid scheduling algorithms. National Weather Services (MET Malaysia) can expand localized wind forecasting for power grid support. MOSTI (Ministry of Science, Technology and Innovation) can fund AI-based research to optimize renewable energy forecasting and balancing. Energy Commission (Suruhanjaya Tenaga) can enforce grid standards that accommodate intermittent renewable sources like wind.

1.2.47 Conclusion

Based on the combined analysis of Figure 6.1, 6.2, 6.3, it is evident that Pulau Pinang consistently experience peak wind actitivity between late morning to afternoon (11AM-4PM), with additional gust surge during early morning (6-8AM), late to midnight (11PM-1AM), reflecting the influence of coastal dynamics and land-sea breeze interactions. In contrast, Kuala Lumpur has a lower and more evenly distributed wind speeds and gusts occurence, only peaking slightly around mid-afternoon (3PM) and during night hours (7-8PM, 1AM). Overall Pulau Pinang displays greater diurnal variability; while Kuala Lumpur remains relatively stable, likely due to urban landscape. This suggested that coastal regions offer more pronounced wind opportunities at specific times of the day, whereas inland urban areas maintain steadier wind conditions, emphasizing the influence of geography on wind variability and the importance of region-specific scheduling for wind energy optimization.

1.3 Report Conclusion

Conclusion

This study has revealed that wind behavior across Malaysia is far from uniform, showing distinct regional and temporal patterns with significant implications for the renewable energy sector. Coastal regions like Penang, particularly areas such as Sungai Puyu and Batu Maung, emerge as strong candidates for wind energy development due to their combination of high mean wind speeds, frequent gust events, and moderate variability. In contrast, inland areas like Sepang demonstrate more stable but lower wind conditions, highlighting the importance of tailored, location-specific energy strategies.

Temporal analysis shows strong seasonal dynamics driven by Malaysia's monsoon cycles. Wind speed and gust intensity increase steadily from April to October, offering a peak window for wind energy capture, while a predictable dip in March to April suggests a need for adaptive energy scheduling and diversified resource portfolios. The observed "U-shaped" long-term trend in average wind speed, from 2018 lows to a moderate post-2021 recovery, underscores the necessity for long-range weather forecasting and multi-year climate modeling in infrastructure planning.

Furthermore, the relationship between wind chill and equipment performance calls for thermal resilience in turbine design, particularly in colder suburban or elevated regions. While wind speed and gust show a strong correlation, other weather variables like humidity and temperature exhibit only weak associations, limiting their predictive value but still offering context for site-specific assessments.

Finally, the identification of regions with frequent extreme wind events, such as Penang Road and Tanjung Bungah, enables prioritization of high-durability infrastructure and real-time forecasting technologies. These findings collectively highlight that successful wind energy deployment in Malaysia requires a nuanced understanding of regional variability, seasonal cycles, and extreme weather risks. Integrating these insights into policy, planning, and engineering design will be key to advancing Malaysia's renewable energy goals in a climate-resilient and economically efficient manner.