

Analysis of Wind Turbine Failure

Business Problem:

Unplanned failure of wind turbine engines is leading to huge losses and negatively impacting electricity generation.

Business Objective: Minimize unplanned failures.

Business Constraint: Maximize power generation.

Success Criteria:

- **Business Success Criteria:** Reduce the unplanned failure of wind turbine engines by at least 30%.
- **Economic Success Criteria:** Achieve a cost saving of at least \$2M per year due to the reduction of unplanned downtime.

Approach: Data Analytics Project Management Methodology.

Metadata for Wind Turbine Dataset:

Variable	Description	Unit
Date	Timestamp of recorded measurement	YYYY-MM-DD
Wind_Speed	Speed of the wind that influences turbine operation	meters per second (m/s)
Power	Power output generated by the wind turbine	kilowatts (kW)
Nacelle_Ambient_Temp	Temperature inside the nacelle	Degrees Celsius (°C)
Generator_Bearing_Temp	Temperature of bearings inside the generator	Degrees Celsius (°C)
Gea_Oil_Temp	Temperature of the gear oil used for lubrication in the gearbox	Degrees Celsius (°C)
Ambient_Temp	External environmental temperature surrounding the wind turbine	Degrees Celsius (°C)
Rotor_Speed	Speed of the rotor blades spinning in response to wind	Revolutions per minute (RPM)
Nacelle_Temp	Overall internal temperature inside the nacelle	Degrees Celsius (°C)
Bearing_Temp	Temperature of the main shaft bearings	Degrees Celsius (°C)
Generator_Speed	Rotational speed of the generator shaft	Revolutions per minute (RPM)
Yaw_Angle	Orientation angle of the wind turbine relative to the wind direction	Degrees (°)
Wind_Direction	The direction from which the wind is blowing	Degrees (°) [0° = North, 90° = East, 180° = South, 270° = West]
Wheel_Hub_Temp	Temperature at the wheel hub which connects the rotor to the main shaft	Degrees Celsius (°C)
Gearbox_Inlet_Temp	Temperature of the gearbox at the oil inlet	Degrees Celsius (°C)
Failure_Status	Indicator of wind turbine failure based on operational conditions	Binary (0 = No Failure, 1 = Failure)

FiMBD: First Moment Business Decision * **SMBD:** Second Moment Business Decision * **TMBD:** Third Moment Business Decision * **FoMBD:** Fourth Moment Business Decision

Q1: First Quartile * **Q3:** Third Quartile * **IQR:** Interquartile Range * **LB:** Lower Bound * **UB:** Upper Bound

Process:

1. Database Creation
2. Data Importing
3. Data Type Casting
4. Rename Columns
5. Duplicates Checking
6. Missing Values Checking
7. EDA
 - First Moment Business Decision
 - Second Moment Business Decision
 - Third Moment Business Decision
 - Fourth Moment Business Decision
8. Outlier Detection – IQR Method
9. Feature Engineering
 - Discretization
10. Scaling

EDA_SQL_Report

Total Features: 16

Total Records Measured: 3600

Timeline: 01-01-2022 to 31-12-2023

Timeline_Status: Inconsistence (Only 2 records of sensor data measured in 2022 – Jan-01 and the remaining 3598 records are meandered from 2023-jan-01 to 2023-Dec-31)

Duplicate Records: 0

Missing Values: 0

Overall Wind Turbine Failures = 33.33%

2022 Annual Total Failures: 1

2022 Annual Failures Rate: 50%

2023 Annual Total Failures: 1199

2023 Annual Failures Rate: 33.32%

	FiMBD			SMBD					TMBD	FoMBD	
Variable	Mean	Median	Mode	Variance	Sta Dev	Min	Max	Range	Skewness	Kurtosis	
Power(kW)	3.313	2.697	1.982	6.816	2.610	0	11.884	11.884	1.602	2.09	
Statistical Insights	1. Mean (3.313 kW) > Median (2.697 kW) indicates positively skewed distribution. 2. Mode (1.982 kW) shows that most common power output (91 times) is on the lower side. 3. Variance (6.816) and Standard Deviation (2.610 kW) indicating moderate spread in the data. 4. Range = 11.884 kW from 0 to peak power output. Indicates significant variability. 5. Skewness = 1.602 Strong right skew most readings are low, but few are very high. 6. Kurtosis = 2.09 High peak and heavy tails and expect more extreme outliers 7. 1037 Outliers found.									Q1	2.21
										Q3	3.24
										IQR	1.02
										LB	0.67
										UB	4.78
Business Insights	1. Mean = 3.313 kW, Mode = 1.982 kW indicating, Turbines frequently operate in low-power zones, likely due to moderate wind conditions or operational constraints.										
	2.Maximum Power = 11.884 kW, Turbines are capable of high power, but such events are rare. Analyze conditions during peak performance to optimize future yield.										
	3. No negative values, Min = 0 Suggests valid readings and proper logging by sensors.										
	4. Power output is fairly stable across time, with less chance of unpredictable spikes or crashes.										
	5. Power generation doesn’t keep pace with available wind during some periods. Investigate turbine calibration or yaw misalignment.										
	6. Wind Turbine failed at maximum Power of 11.88 kW on 5 th Feb Aug 2023.										
	7. Wind Turbine not failed at Power generation of 6.96 kW, which recorded on 29 th Dec 2023. So, the limit of power generation is 6.96 kW, beyond that turbine will fail.										
	8. Wind Turbine failed at maximum wind speed of 44.51 m/s on 8 th Feb Oct 2023 and generated power of 0.59kW. But at the wind speed of 4.22 m/s, the power generated is 6.96 Kw, which is threshold but the turbine is not failed at those readings. Need to verify the seasonality trends and sensor calibration.										

	FiMBD			SMBD					TMBD	FoMBD	
Variable	Mean	Median	Mode	Variance	Sta Dev	Min	Max	Range	Skewness	Kurtosis	
Generator_Bearing_Temp(°C)	84.07	83.123	85.596	800.2	28.287	9.91	143.56	133.64	-0.127	0.39	
Statistical Insights	1. Mean ≈ Median ≈ Mode Distribution appears to be approximately symmetric, with no extreme skew. 2. Standard Deviation (28.28) and Range (133.64°C) There’s a wide temperature spread - needs domain validation to check acceptable thermal boundaries. 3. Low Skewness (-0.127) Slight left skew, but nearly symmetrical distribution -good for statistical modeling. 4. Kurtosis 0.39 Indicates Slightly more peaked than normal, but not alarming. Mild outliers might be present. 5. Min (9.91°C) seems too low Generator bearings should rarely operate below ambient – this could be a sensor fault or downtime reading. 6. 637 Outliers found.									Q1	72.58
										Q3	94.96
										IQR	22.375
										LB	39.023
										UB	128.52
Business Insights	1. Max temp recorded is 143.56°C, which might exceed the safe operating limits of generator bearings. Consistently high temperatures reduce bearing lifespan and can trigger turbine shutdowns.										
	2. Minimum of 9.91°C is abnormally low - likely from: <ul style="list-style-type: none">• Sensor misreading• Periods of inactivity (e.g., shutdown or maintenance)										
	3. High Kurtosis (123.7) means sharp peaks in data, suggests transient overheating events that need real-time alerts.										
	4. Wind turbine failed at a minimum Generator_Bearing_Temp of 9.91°C on 10 th Mar 2023 and failed at a maximum Generator_Bearing_Temp of 143.56°C on 13 th Sep 2023.										
	5. Wind Turbine not failed at Generator_Bearing_Temp of 98.98°C, which recorded on 16 th Jun 2023. So, the limit of Generator_Bearing_Temp of 98.98°C, beyond that turbine will fail.										

	FiMBD			SMBD					TMBD	FoMBD	
Variable	Mean	Median	Mode	Variance	Sta Dev	Min	Max	Range	Skewness	Kurtosis	
Gear_Oil_Temp(°C)	84.278	76.972	89.127	529.410	23	49.515	128.723	79.208	0.431	-1.13	
Statistical Insights	1. Mean > Median Indicates right-skewed distribution, though Skewness (0.431) is mild — most values slightly lean toward higher temps. 2. Significant spread in gear oil temps, suggesting variability in operating loads or environmental conditions. 3. Wide Range (≈79°C) A large operating window, indicating potential load changes, turbine idling, or operational stress. 4. Kurtosis is -1.13, Very flat distribution. Data is more uniformly distributed with very few extreme values. 5. No outliers found.									Q1	65.96
										Q3	106.37
										IQR	40.40
										LB	5.34
										UB	166.98
Business Insights	1. Max gear oil temperature recorded: 128.72 °C Compare with OEM-specified maximum (typically 110–120 °C). If exceeded, initiate: <ul style="list-style-type: none">Gearbox inspectionLubricant condition assessmentCooling system validation										
	2. Higher average and peak temperatures can accelerate oil degradation. <ul style="list-style-type: none">Implement a predictive maintenance schedule based on gear oil thermal exposurePerform oil viscosity, TAN (Total Acid Number) tests periodically										

