Preprocessing Methodology:

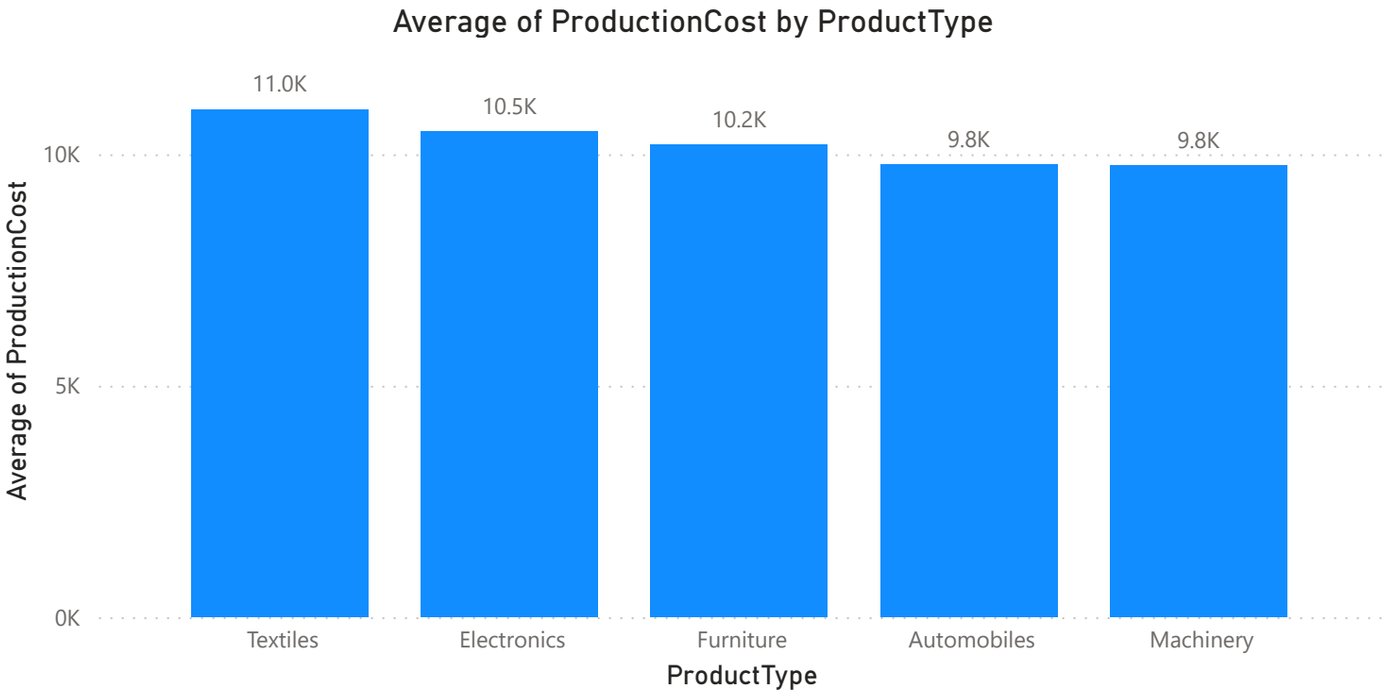
- As a first step in this project, we established a connection with Excel, After that we have imported the two datasets and renamed them as "**product_data**" and "**employee_data.**" We moved to the model view , there we observed that Power BI automatically detected a relationship between the datasets based on the "**ProductID**" column.
- Now moving to the **Power Query Editor**, we identified missing values in "**ProductionCost**" and "**QuantityProduced**" in "**product_data**" and in "**Salary**" and "**PerformanceRating**" in "**employee_data.**" Additionally, we noted irrelevant data in the "**Warehouselocation"** and "**Department**" columns.
- Now to address missing "**Department**" values in "**employee_data**", we utilized "**Employee Training Records**" and replaced irrelevant data with the "**Logistics**" department.
- For missing "ProductionCost" and "QuantityProduced" values in "product_data," we employed a strategy of **mean** and **median**, respectively, based on "ProductType" and "CountryofOrigin." We created a **groupby** table and **merged** it with the original data after handling null values.
- Similarly, for "Salary" and "PerformanceRating" in "employee_data," we applied **mean** and **median** strategies based on "Department" and "CountryofOperations," creating a groupby table, handling null values, and merging it back with the original data.
- After filling missing values, we merged "product_data" and "employee_data" on the "ProductID" column, creating a final merged dataset named '**Final Data**.'
- At the end of preprocessing, we have dropped the rows which contains the irrelevant data in the "WarehouesLocation" column.
- We have imported the 'Final Data' into the Power BI from the Power query Editor.

Analysis of Production Costs

- Calculate the average production cost for each product type. Which product type has the highest average cost?

Average Cost By Product Type

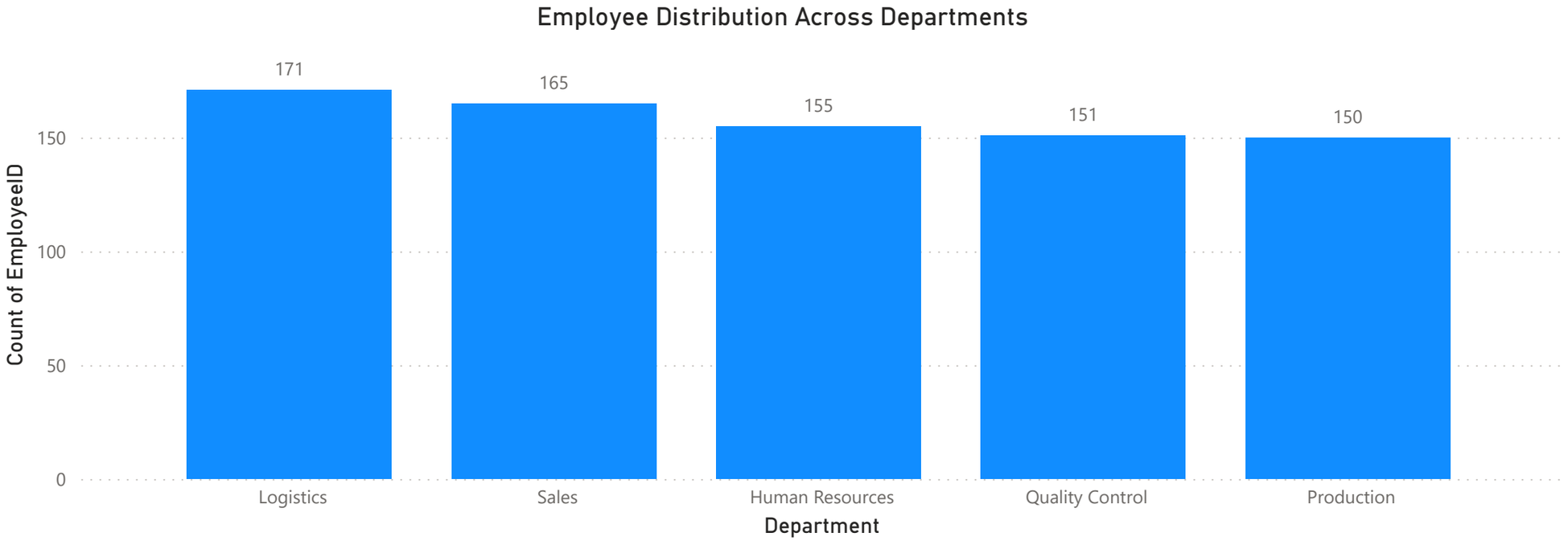
ProductType	Average of ProductionCost
Automobiles	9,773.85
Electronics	10,494.53
Furniture	10,205.17
Machinery	9,755.21
Textiles	10,966.43
Total	**10,260.83**



From the graph, it is evident that Textile products have higher average production cost compared to other product types.

Employee Distribution Across Departments

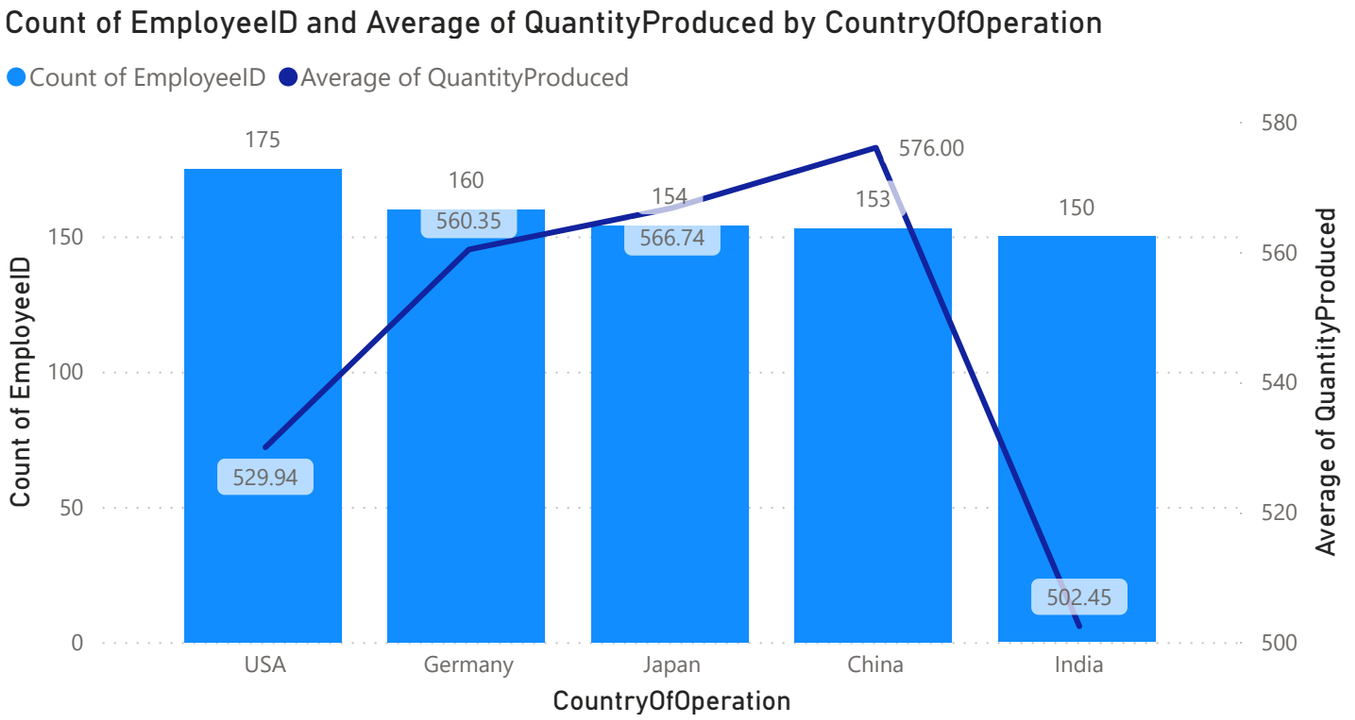
- Analyze the distribution of employees across different departments. Which department has the most employees?



Logistics department has the most number of employees compared to other departments.

Country-Based Analysis of Operations

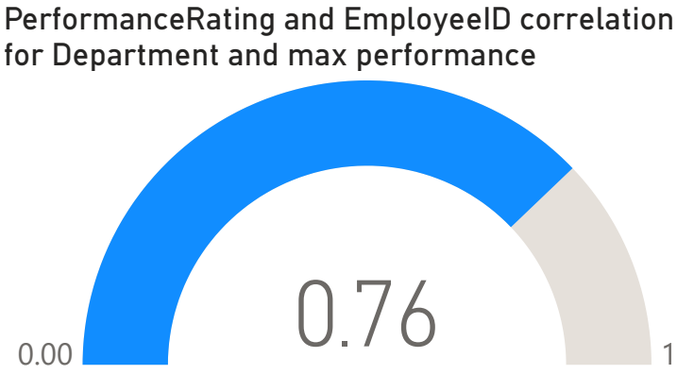
- Investigate which country has the highest number of employees and the highest average production.

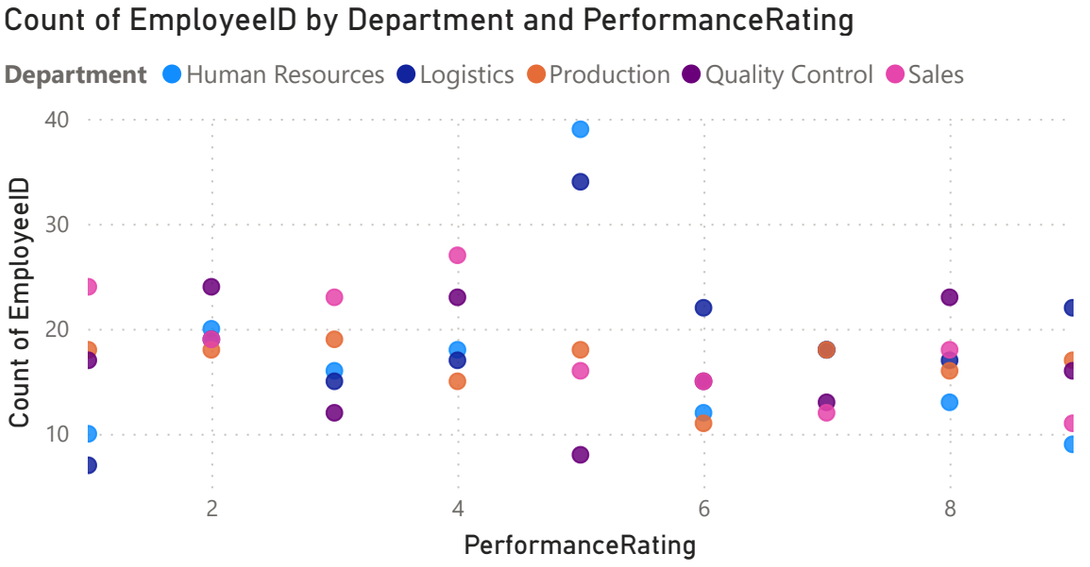


From the graph it is evident that USA has more number of employees compared to other countries whereas China has the highest average production. Here in this question the secondary y-axis is production cost it is better quantity produced.

9.Performance Rating Analysis

- Using DAX, analyze the average performance rating by department. Is there a correlation between department and performance rating?



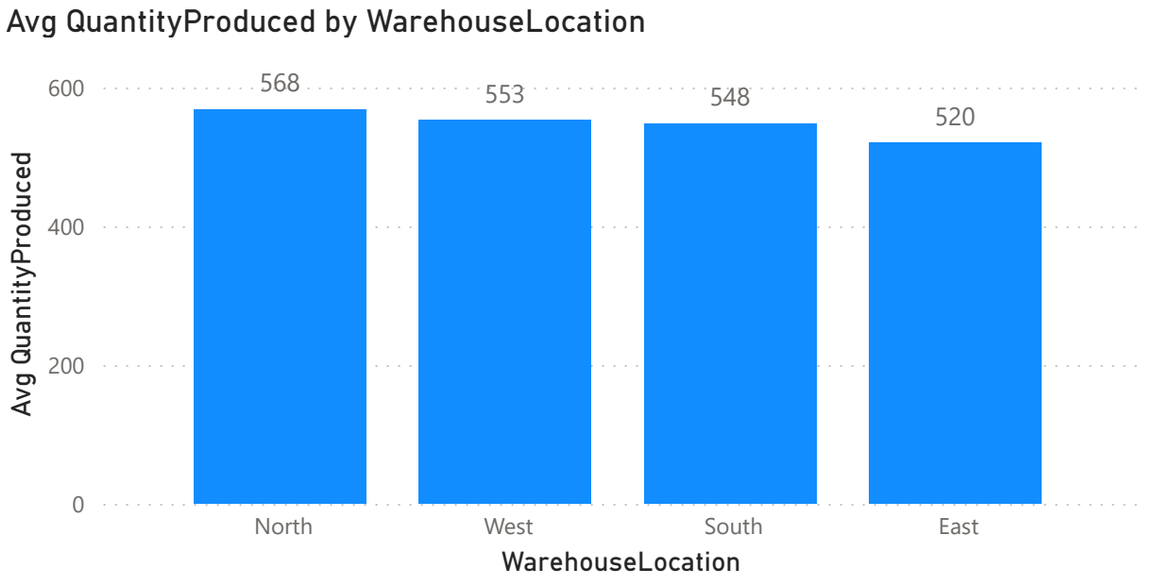


Observation:

Based on the gauge plot obtained , we can observe that the Department and performance rating is Positively correlated with a correlation coefficient of 0.76

Warehouse Efficiency Analysis

- Calculate the average quantity of products stored in each warehouse location. Which warehouse location is utilized the most?

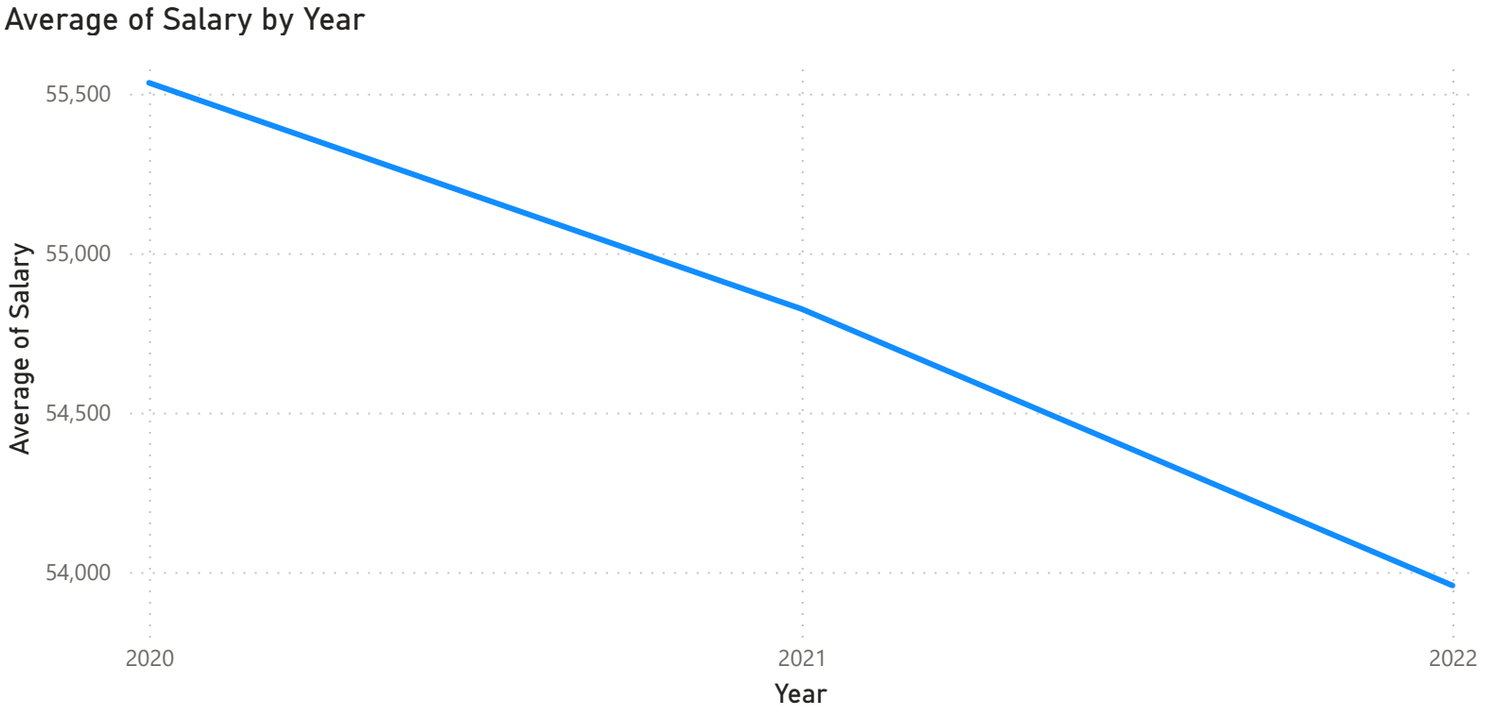


Observations

- Ware house in the North location has produced more quantity compared to other warehouses.

Salary Trends Over Time

- Analyze the trends in salaries over time. Are there noticeable increases or disparities?



30,013.47

Min of Salary

79,990.83

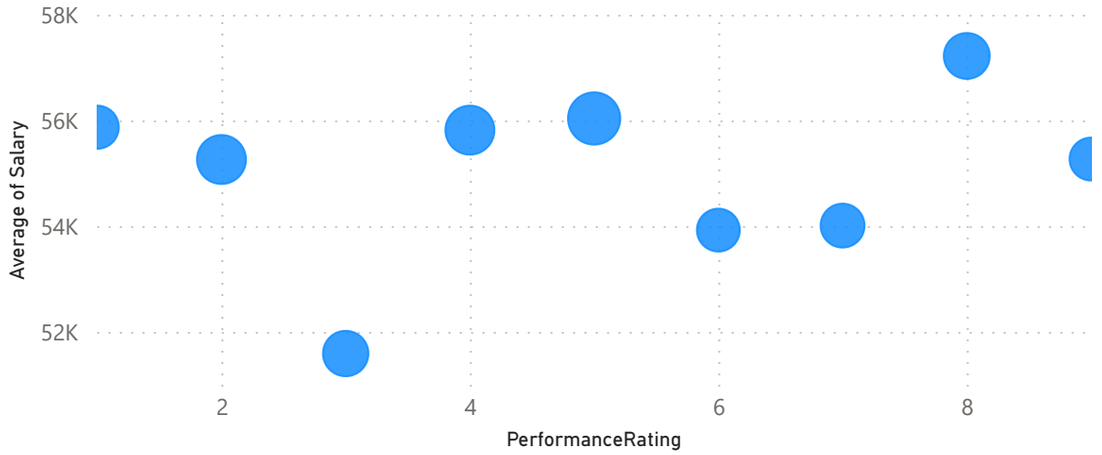
Max of Salary

Observations:

- From the graph it is clear that there is decline in the average salary of the employees over the years.

12. Correlation Between Salary and Performance

- Explore if there's a correlation between employees' salaries and their performance ratings.



Correlation coef by Department

0.46

Correlation coef by countryOfOperstion

-0.49

Correlation coef by countryOfOrigin

0.45

Observation:

- To explore the correlation between employee's salaries and their performance ratings it is recommended to use the quick measure ,an in-built function in Power BI. Based on the above methodology, it is observed that :
- Based on Department, the Average salary and Performance rating are **positively correlated** with a correlation factor of **0.46**.
- Based on Country of Origin, The Average salary and Performance rating are **positively correlated** with a correlation factor of **0.45**.
- Based on Country of Operation, The Average salary and Performance rating are **negatively correlated** with a correlation factor of **0.49**.

Product Manufacturing Trends

Analyze how the manufacturing of different product types has trended over time. Are there seasonal patterns?

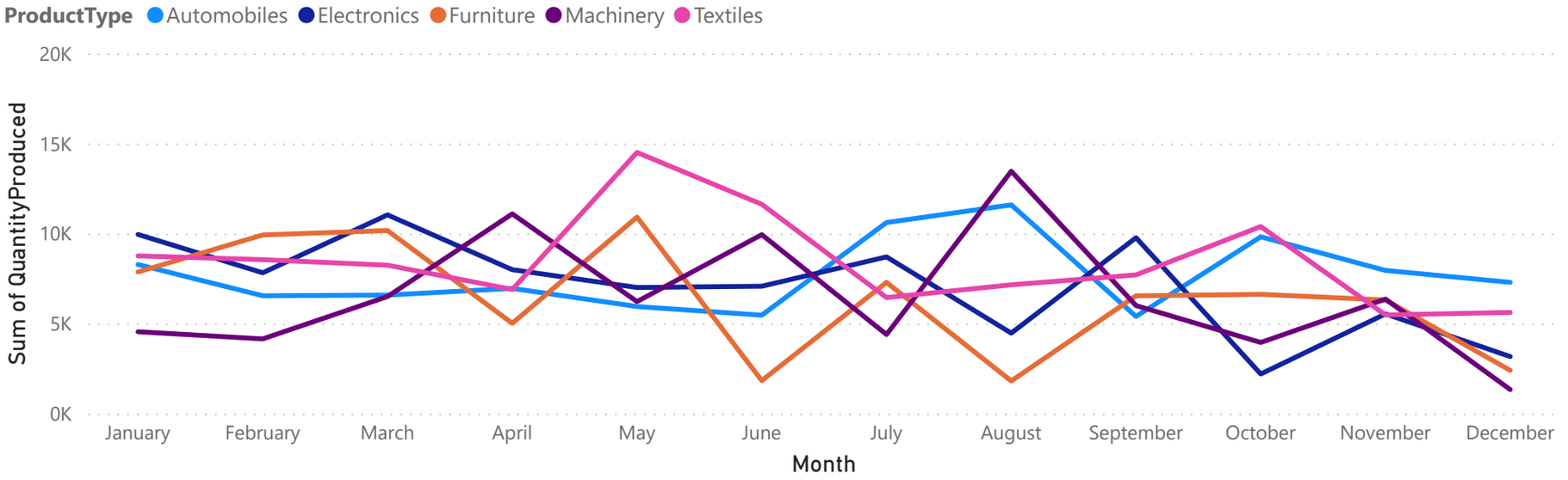
Year

2023 | 2024 | 2025

ProductType

- [ ] Automobiles
- [ ] Electronics
- [ ] Furniture
- [ ] Machinery
- [ ] Textiles

Sum of QuantityProduced by Month and ProductType

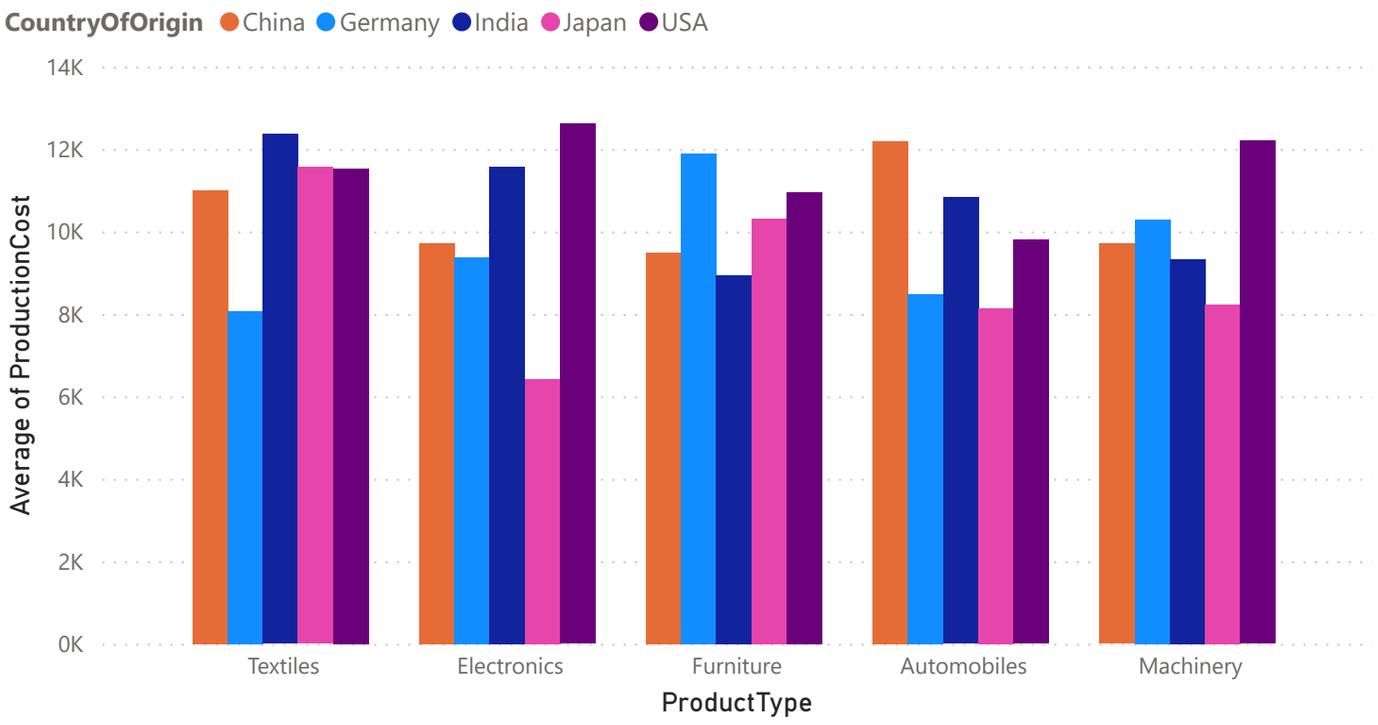


Observations

- We observe an increasing trend in Automobile from June.
- We observe No trend for electronics till July ,followed by a slight decreasing trend till December.
- We observe an increasing trend for Furniture till march, followed by a decreasing trend till august followed by an increasing trend till December.
- We observe an increasing trend in Machinery till march ,followed by almost no trend till September and a decreasing trend till dec.
- For Textile we observe almost no trend throughout except for the month of April where the production increases and peaks in May.

Cost Analysis by Country of Origin

- Investigate the average production cost per product in each country of origin. Which country has th highest and lowest costs?



ProductType	China	Germany	India	Japan	USA	**Total**
Machinery	9,706.83	10,294.58	9,325.41	8,234.71	12,200.47	**9,755.21**
Automobiles	12,191.38	8,480.09	10,829.88	8,141.09	9,800.20	**9,773.85**
Furniture	9,495.63	11,896.56	8,944.94	10,318.38	10,949.42	**10,205.17**
Electronics	9,710.41	9,377.58	11,566.27	6,420.49	12,612.91	**10,494.53**
Textiles	10,996.46	8,075.05	12,364.78	11,562.27	11,532.26	**10,966.43**
Total	**10,501.57**	**9,492.62**	**10,717.83**	**9,236.31**	**11,375.34**	**10,260.83**

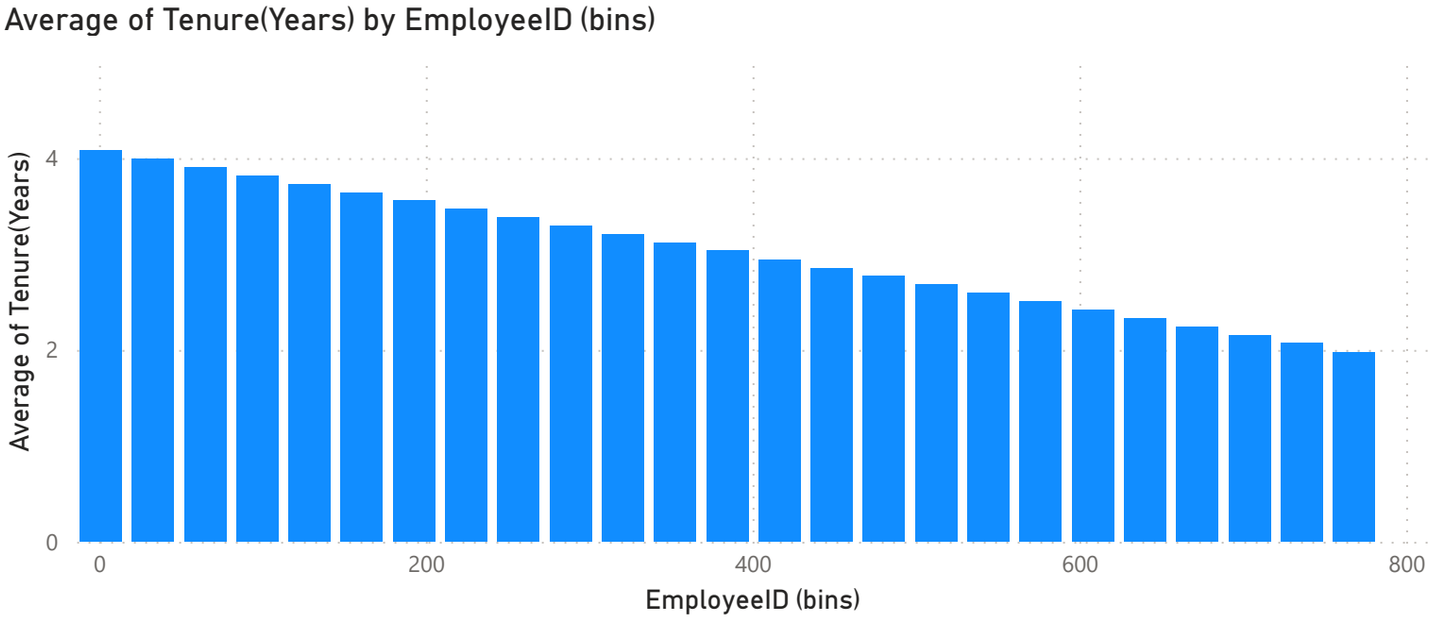
Conclusion :

- To Investigate the average production cost per product in each country of origin, it is recommended to used stacked column graph for ease of visual inspection. Based on the stacked column graph obtained ,it can be concluded that:

Product type	Highest Average Cost	Lowest Average Cost
Textile	India	Germany
Electronics	USA	Japan
Furniture	Germany	India
Automobiles	China	Japan
Machinery	USA	Japan

Employee Tenure Analysis

- Calculate the tenure of employees in the company and analyze its distribution.



3.0

Average of Tenure(Years)

Conclusion:

- To calculate the tenure of employees in the company and analyze its distribution, Line and cluster column with bins graph is suggested . Based on the graph obtained , it can be concluded that the Employees in the company have varied range of experience with maximum being 4 years and minimum being 1.9 years. The average tenure(in years) of the employees is 2.9 years

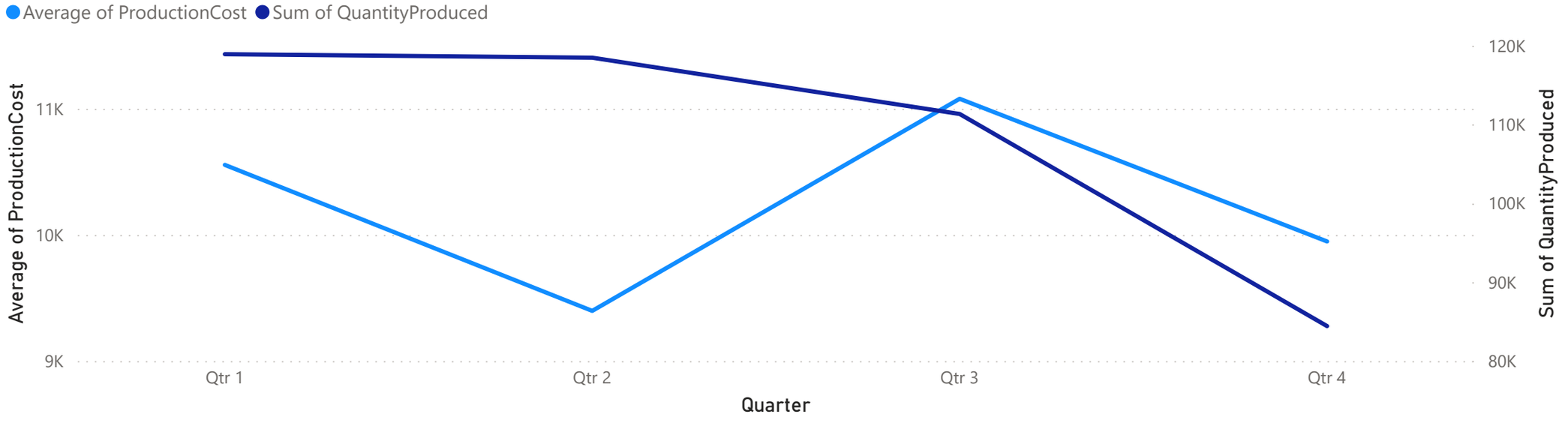
Sum of QuantityProduced and Count of ProductID by PerformanceRating

CountryOfOrigin: All

ProductType: All

ProductType	Count of ProductID	Sum of QuantityProduced	Average of ProductionCost
Automobiles	172	92439	9,773.85
Electronics	156	84727	10,494.53
Furniture	142	76693	10,205.17
Machinery	142	77946	9,755.21
Textiles	180	101364	10,966.43
Total	**792**	**433169**	**10,260.83**

Average of ProductionCost and Sum of QuantityProduced by Quarter



Observations:

- As the Product Lifecycle will represent the introducing the products into the market, growth of the products in the market, reaching the Peak sales in the market and decline of the product sales. we have used the production quantity and Production time line to represent the Product Life cycle.
- The line plot reveals a correlation between production quantity and production cost, indicating a decrease in production quantity when production costs are high. Notably, manufacturing in bulk during the 1st and 2nd quarters results in higher production quantities and lower production costs. However, a declining trend in production quantity is evident from the 3rd quarter onward.
- Examining the scatter plot, it highlights the influence of employee performance on production quantity. Notably, employees with a performance rating of 5 positively impact production quantity.

Advanced DAX: Cost Efficiency Analysis

- Using DAX, explore the cost efficiency of production (production cost per unit of product).

Cost per unit Product

ProductID	Automobiles	Electronics	Furniture	Machinery	Textiles	**Total**
2			25.87			**25.87**
4		28.11				**28.11**
5		75.73				**75.73**
6					52.09	**52.09**
8					15.41	**15.41**
Total	**4115.84**	**4091.63**	**4060.99**	**3732.79**	**5078.33**	**21079.58**

Total Quantity Produced

ProductID	Automobiles	Electronics	Furniture	Machinery	Textiles	**Total**
2			513			**513**
4		384				**384**
5		4815				**4815**
6					1506	**1506**
8					1462	**1462**
Total	**92439**	**84727**	**76693**	**77946**	**101364**	**433169**

Total Production cost

ProductID	Automobiles	Electronics	Furniture	Machinery	Textiles	**Total**
2			13,271.45			**13,271.45**
4		10,795.60				**10,795.60**
5		72,931.04				**72,931.04**
6					39,221.14	**39,221.14**
8					11,268.11	**11,268.11**
Total	**16,81,102.17**	**16,37,147.43**	**14,49,133.82**	**13,85,239.47**	**19,73,958.08**	**81,26,580.97**

Extracting Key Information

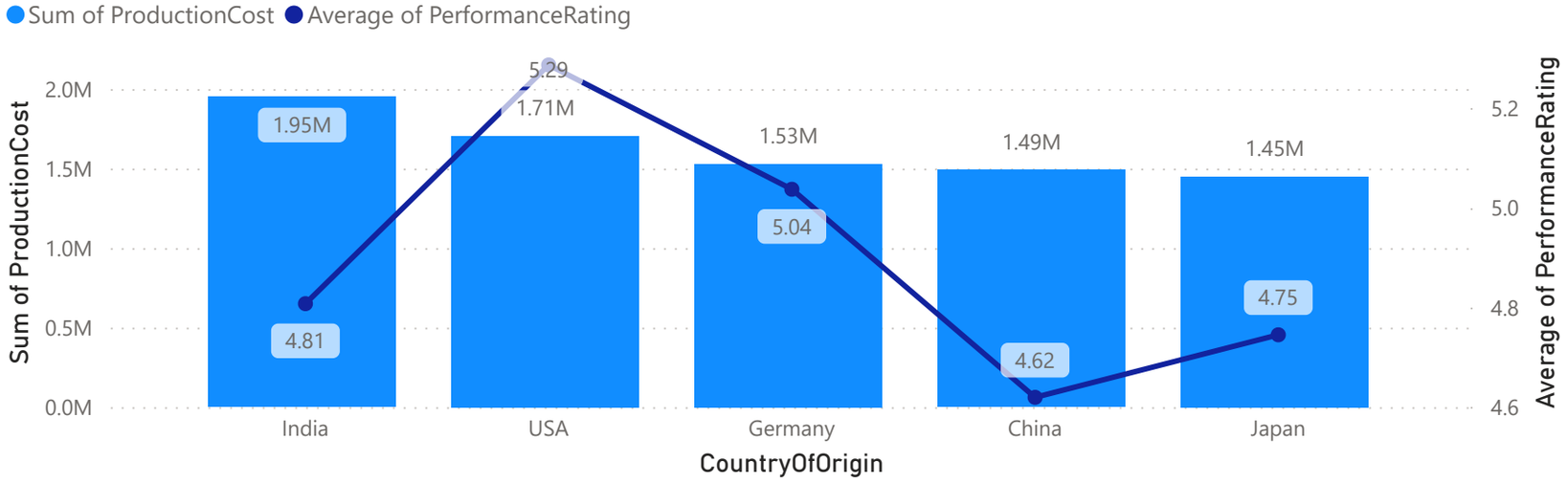
"Using the 'Employee Training Record' column in the Manufacturing Dataset 2, create two new columns. One column should list the dates of all training sessions for each employee, and the other should list the types of training sessions (e.g., Sales Techniques Workshop, Leadership Skills Seminar)."

- Firstly, split the column using split option using delimiters ':' ,';'. The column is split into 6 columns with 3 columns representing the training dates and remaining 3 columns for training the employee attended.
- Every employee did at least one training.
- Then merged all the date columns with separator as ',' and renamed.
- Also merged all the training name columns with ',' as separator and renamed.
- By following these steps key information is extracted like dates list, trainings list from the 'Employee Training Record'

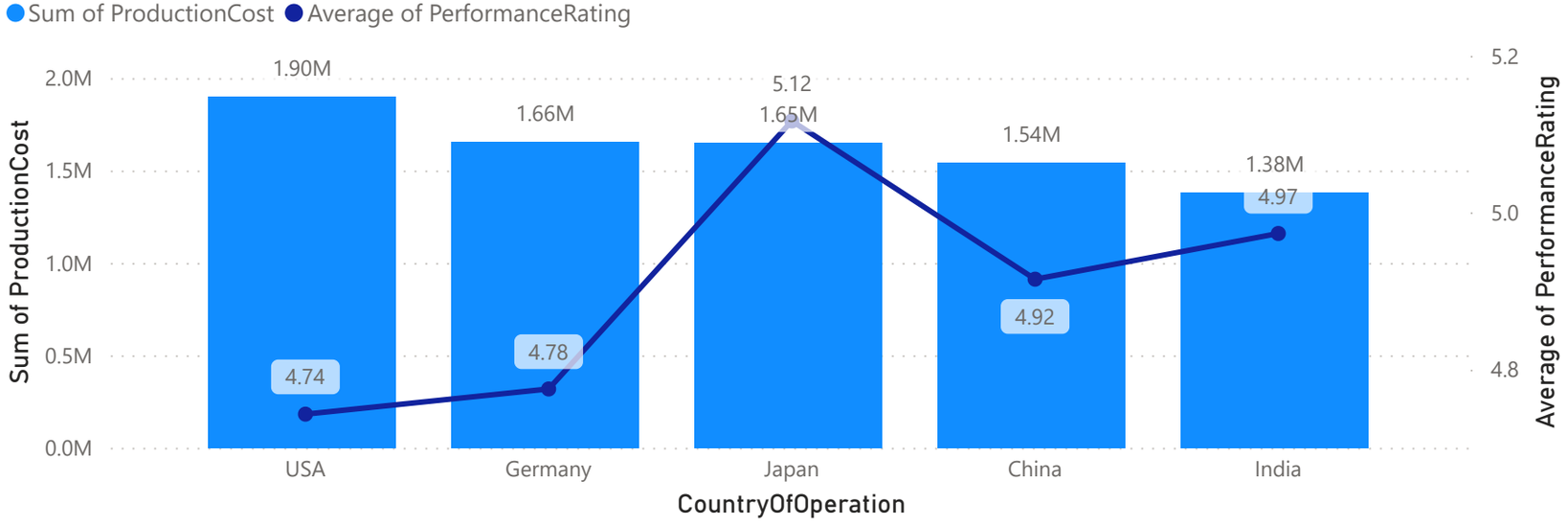
Country of Operation vs. Country of Origin

Compare the countries of operation and origin in terms of production and employee performance.

Sum of ProductionCost and Average of PerformanceRating by CountryOfOrigin



Sum of ProductionCost and Average of PerformanceRating by CountryOfOperation

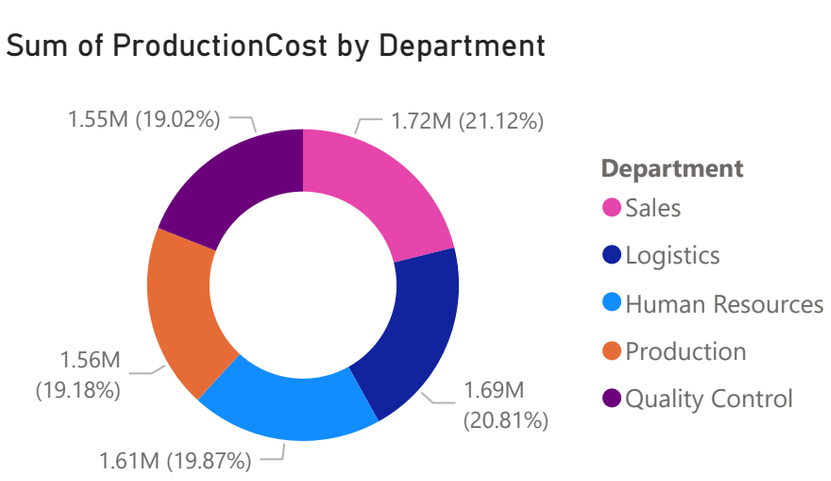


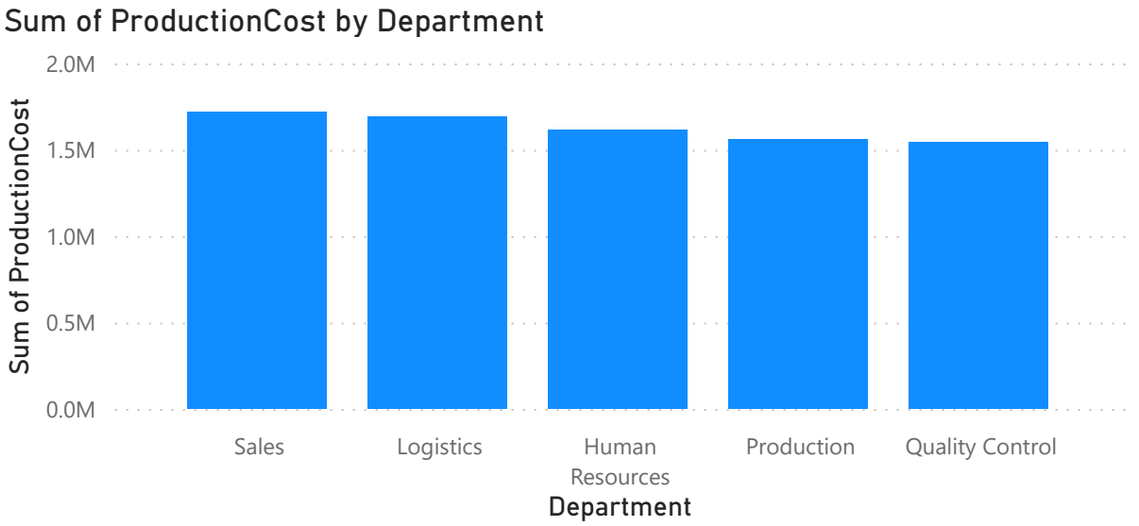
Observation:

1. Production cost originally from India is greater at 1.95 million, with the United States, Germany, and China having the lowest costs. On the other hand, the USA has the highest performance rating, followed by the United States, Germany, India, Japan, and China.
2. The operational production cost from the United States is 5.54, greater than that of Germany, Japan, China, and India. India has the highest performance rating, followed by the United States, China, Japan, and Germany.

Employee Role in Production Cost

Analyze if certain departments or employee roles have a significant impact on production costs.





Department	Sum of ProductionCost	Count of EmployeeID
Human Resources	16,14,790.09	155
Logistics	16,91,008.68	171
Production	15,59,069.08	150
Quality Control	15,45,440.37	151
Sales	17,16,272.75	165
Total	**81,26,580.97**	**792**

Observations:

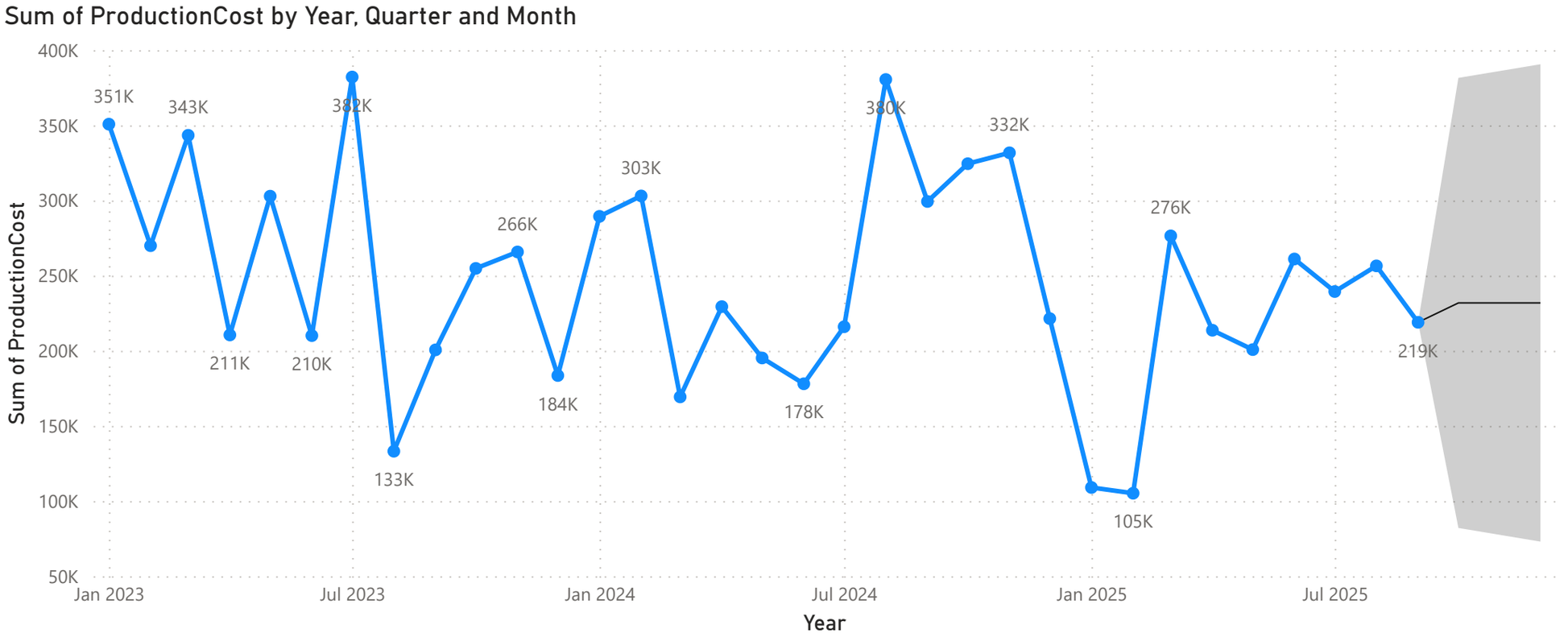
Employee Role in Production Cost:

The employee with the highest production cost is in the sales department, while the lowest is in the quality control department.

The sales department accounts for 21.2% of the overall cost of production, followed by logistics (20.81%), human resources (19.87%), production (19.18%), and quality control (19.02%).

Data Modeling: Time Series Forecasting of Costs

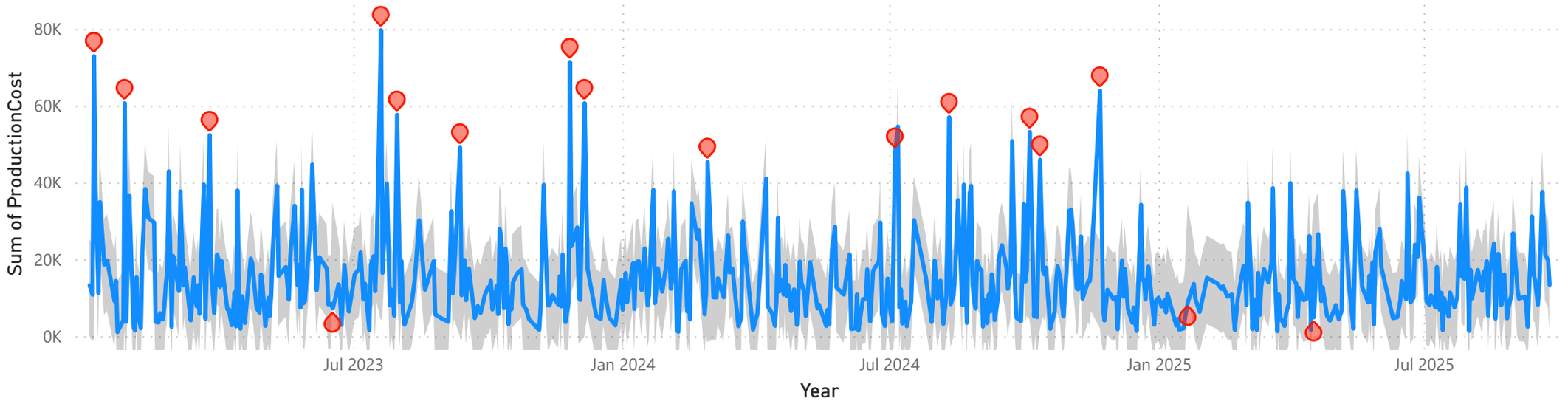
- Perform time series forecasting of production costs using historical data. What are the predicted costs for the next quarter?



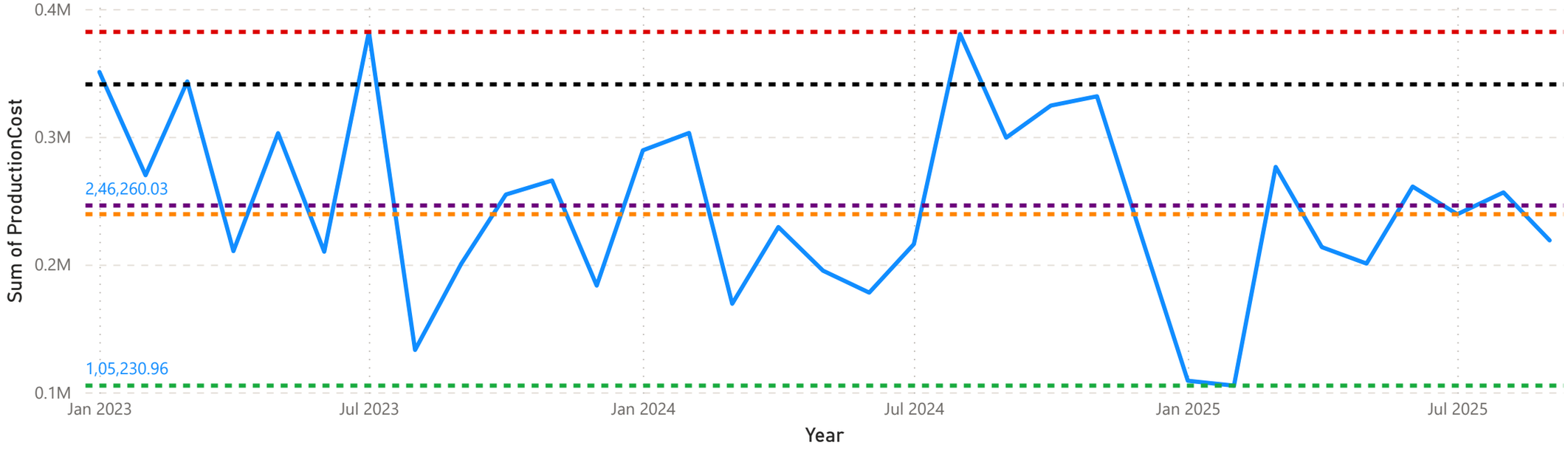
Conclusion:

- In Time Series Forecasting of Costs, We have predicted the manufacturing cost for the upcoming quarter using analytics, and that cost is expected to be 231842.30 with an upper bound of 390,638.78 and a lower bound of 73045.83.

Sum of ProductionCost by Year, Quarter, Month and Day

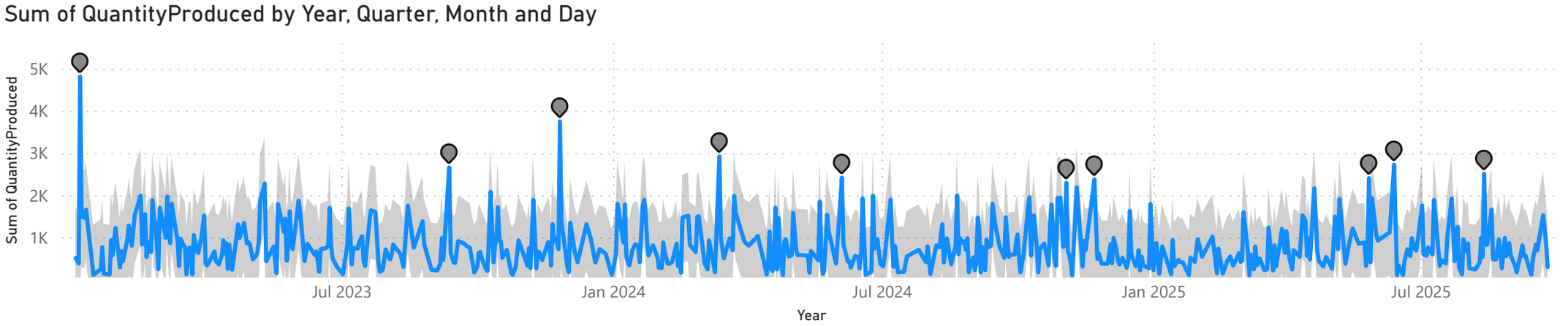


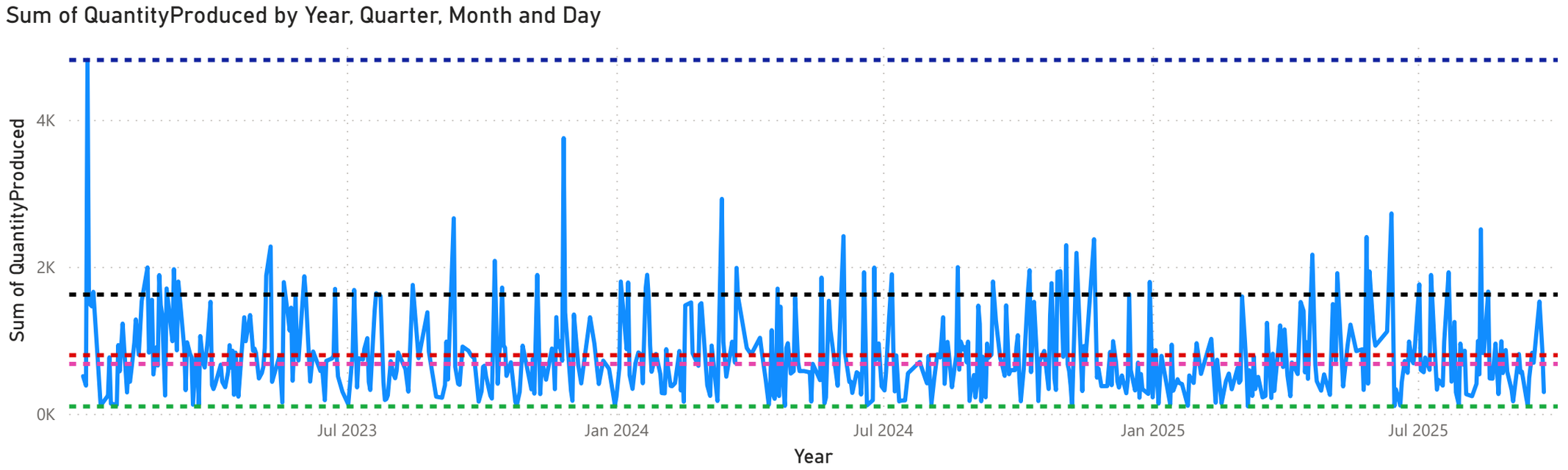
Sum of ProductionCost by Year, Quarter and Month



Advanced Data Transformation: Identifying Production Anomalies

-
- Using Power BI's data transformation capabilities, identify any anomalies in production data (e.g., unusually high costs, sudden spikes in production quantity).





Conclusion:

We can infer from the graph that both the quantity produced and the production cost exhibit anomalies 10 anomalies in the quantity produced column and eighteen in the production cost column.

Manufacturing Report

warehouse Location: East | North | South

Production Year: 2023 | 2024 | 2025

Department: All

Product Type: All

Country: All

ProductType	China	Germany	India	Japan	USA	Total
Automobiles	892.15	844.33	844.08	862.58	672.69	**4115.84**
Electronics	334.15	873.09	1425.20	210.55	1248.64	**4091.63**
Furniture	824.81	739.24	828.03	514.45	1154.47	**4060.99**
Machinery	941.30	831.61	962.02	529.79	468.08	**3732.79**
Textiles	915.22	771.28	1520.80	1168.11	702.92	**5078.33**
Total	**3907.63**	**4059.56**	**5580.12**	**3285.48**	**4246.79**	**21079.58**

Total Production Cost

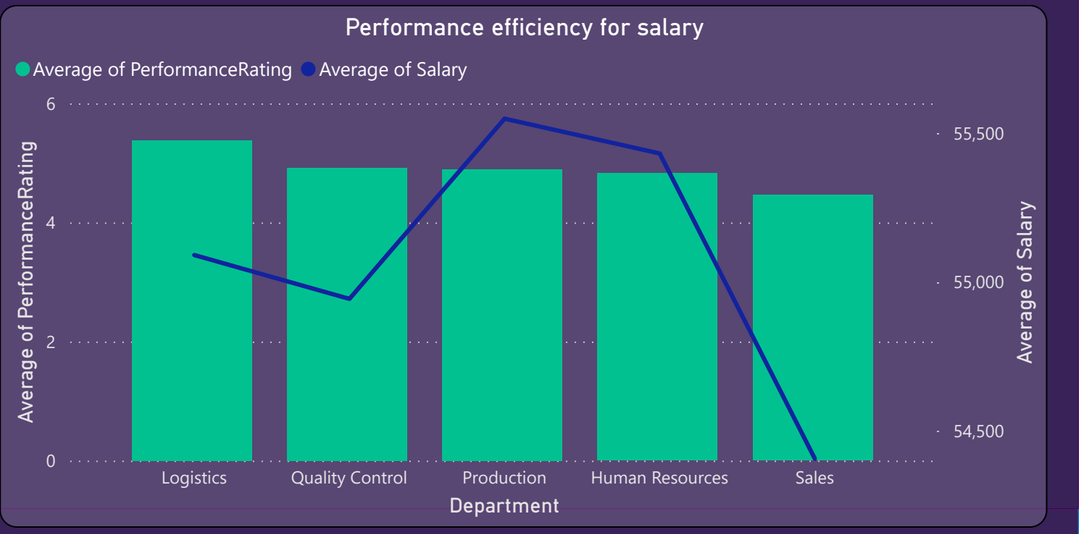
8.13M

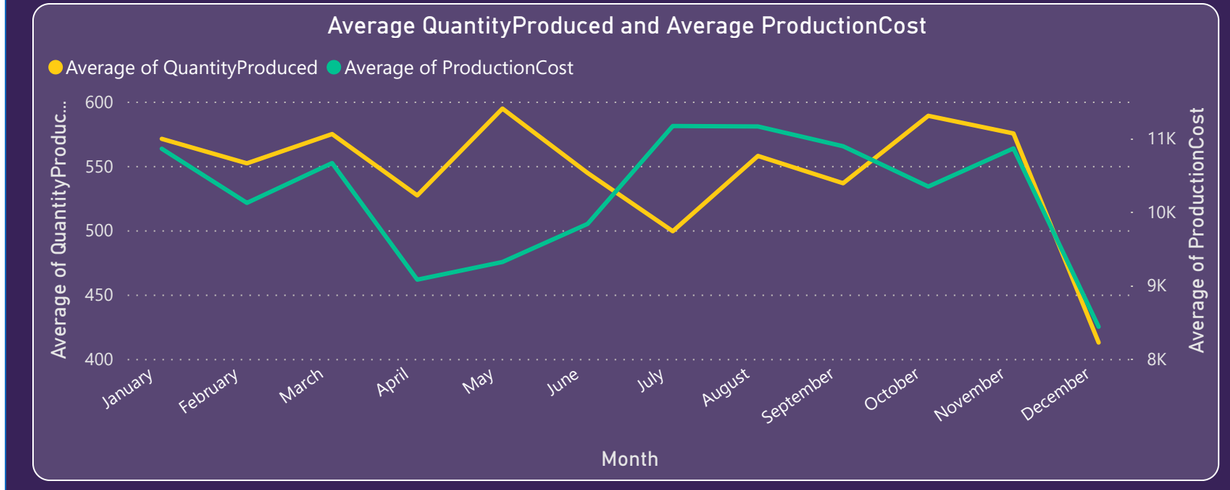
Average Salary

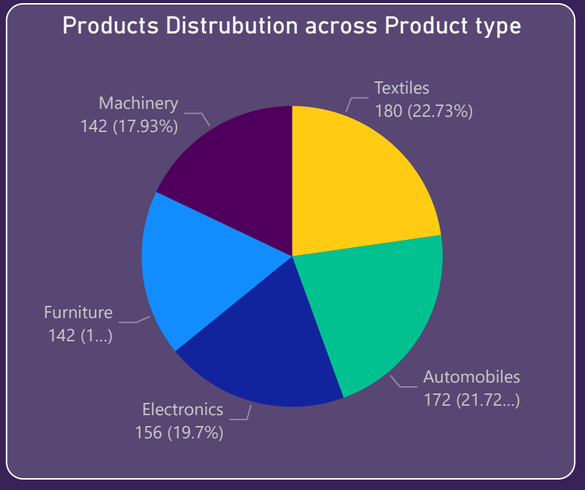
55.07K

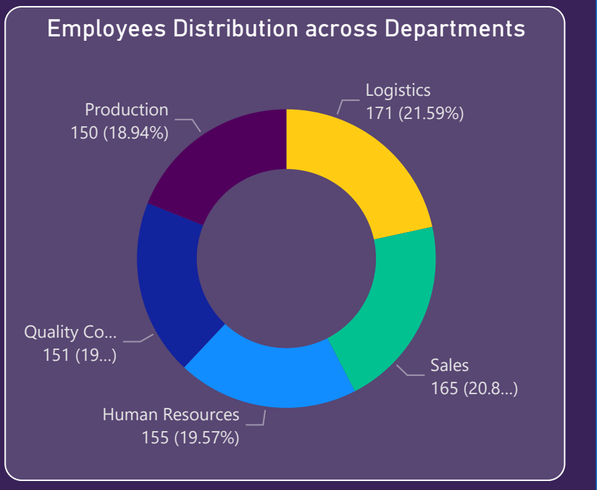
Total Products

792









Time-Based Analysis of Salary and Production trends

warehouse Location: East | North

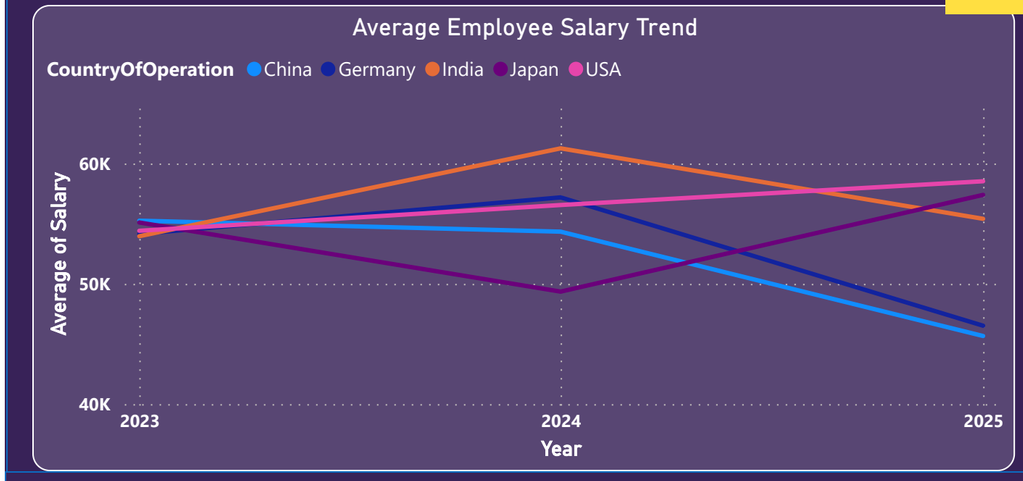
Production Year: 2023 | 2024 | 2025

Hiring Year: 2020 | 2021 | 2022

Department: All

Product Type: All

Country: All



Average Salary

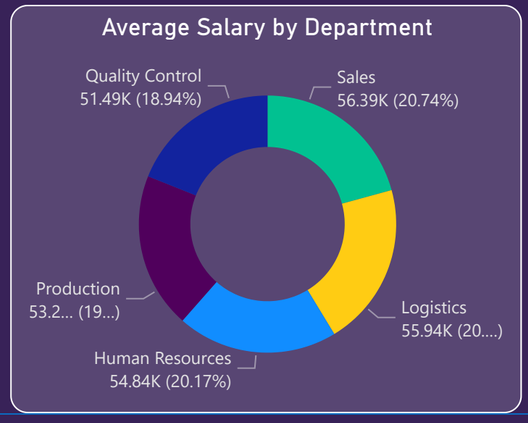
54.34K

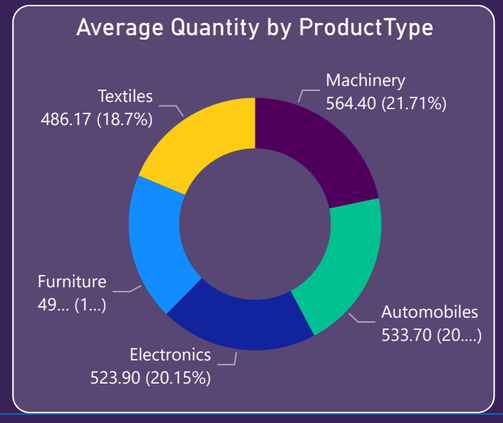
Average Production Cost

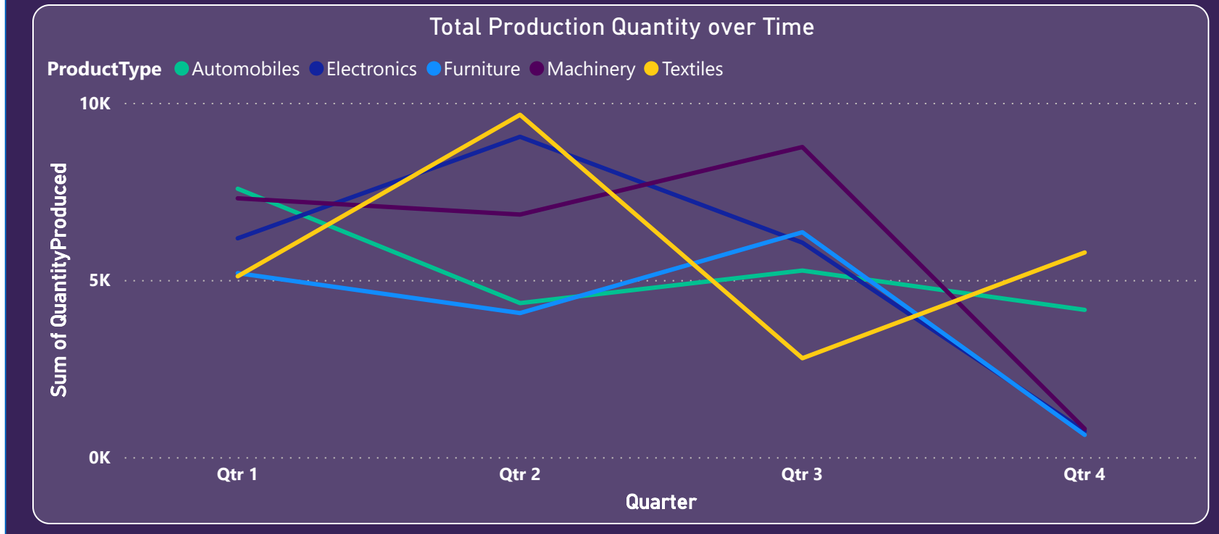
9.65K

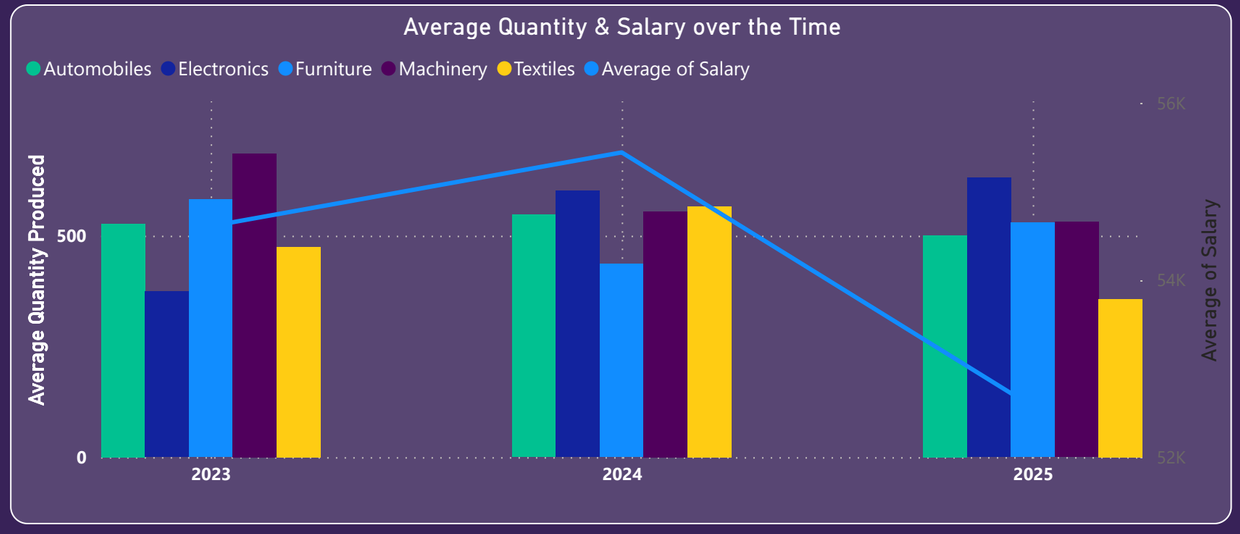
Average Production Quantity

520.15









Analysis Of Employee Data

Hiring Year

2020 | 2021 | 2022

Country of Operations

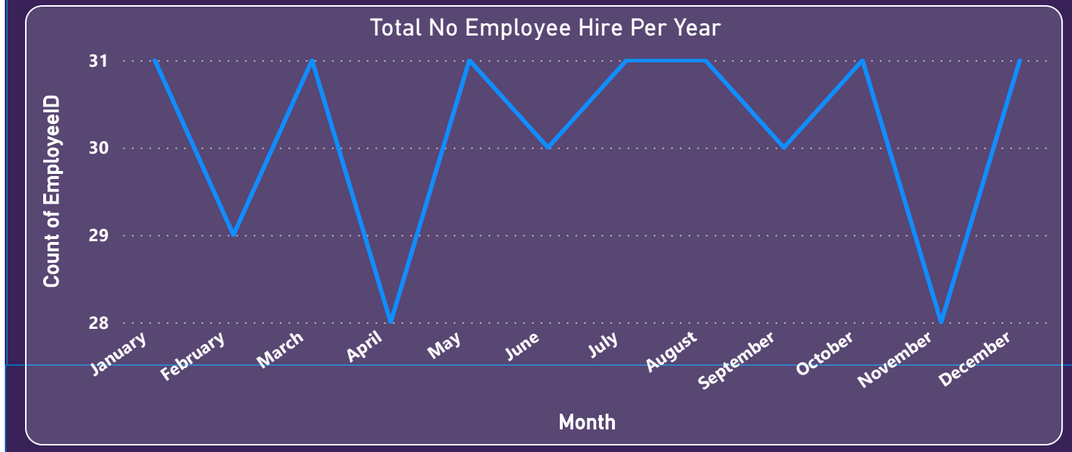
All

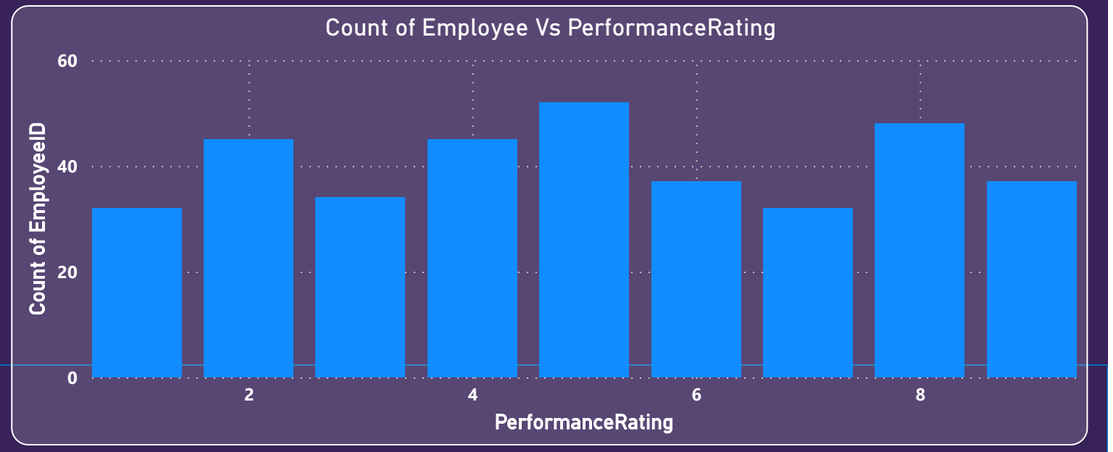
Total Salary

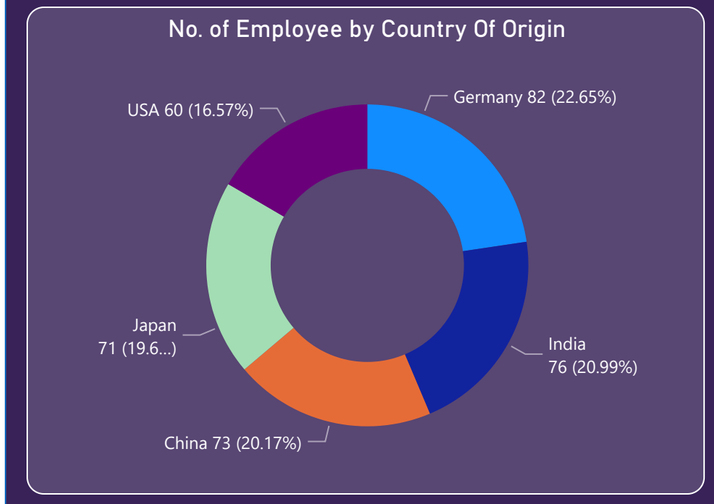
20.10M

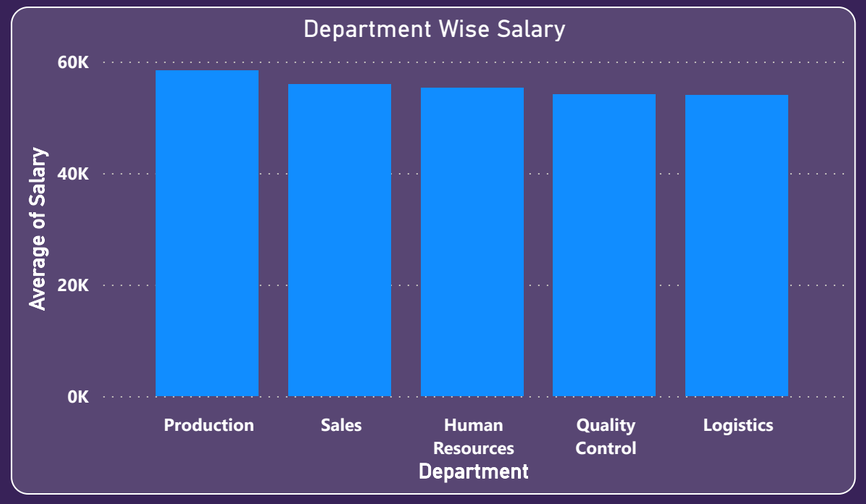
No. of Employees

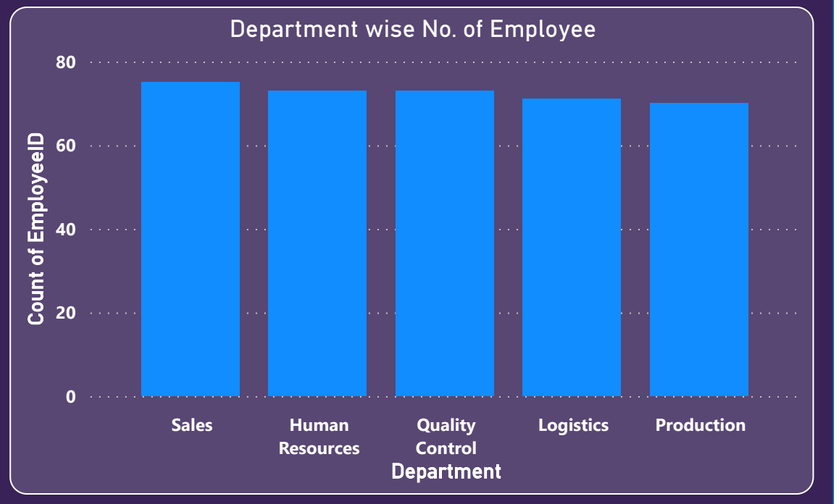
362











warehouse Location

East | North

Production Year

2023 | 2024 | 2025

Hiring Year

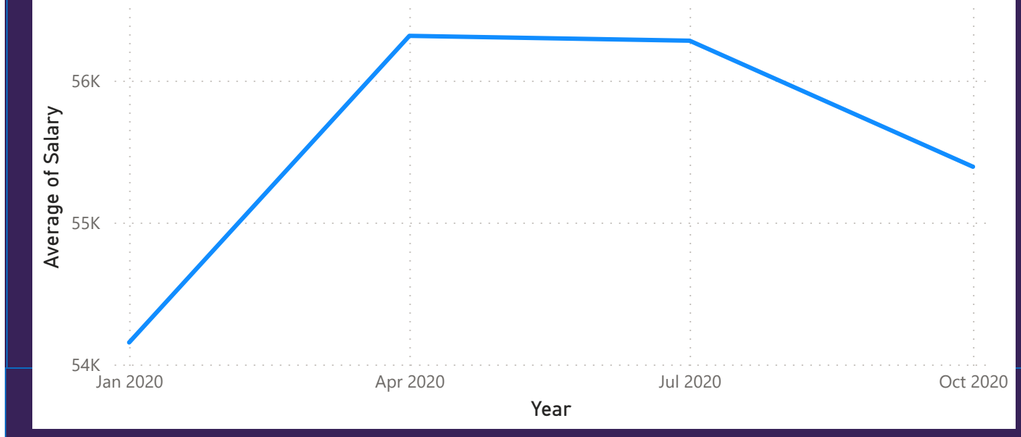
2020 | 2021 | 2022

Department: All

Product Type: All

Country: All

Average Employee Salary Trend



Average Salary

55.53K

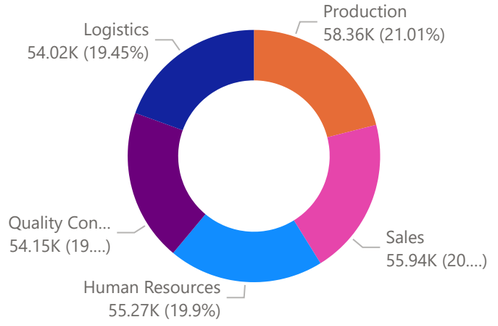
Average Production Cost

10.17K

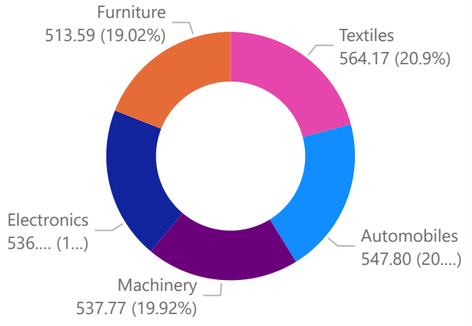
Average Production Quantity

540.88

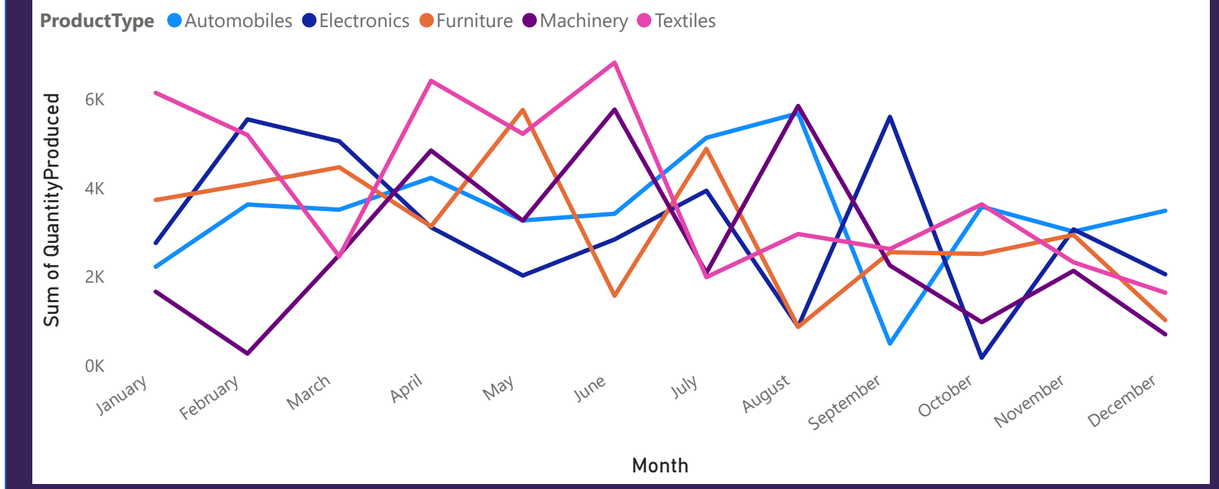
Average Salary by Department



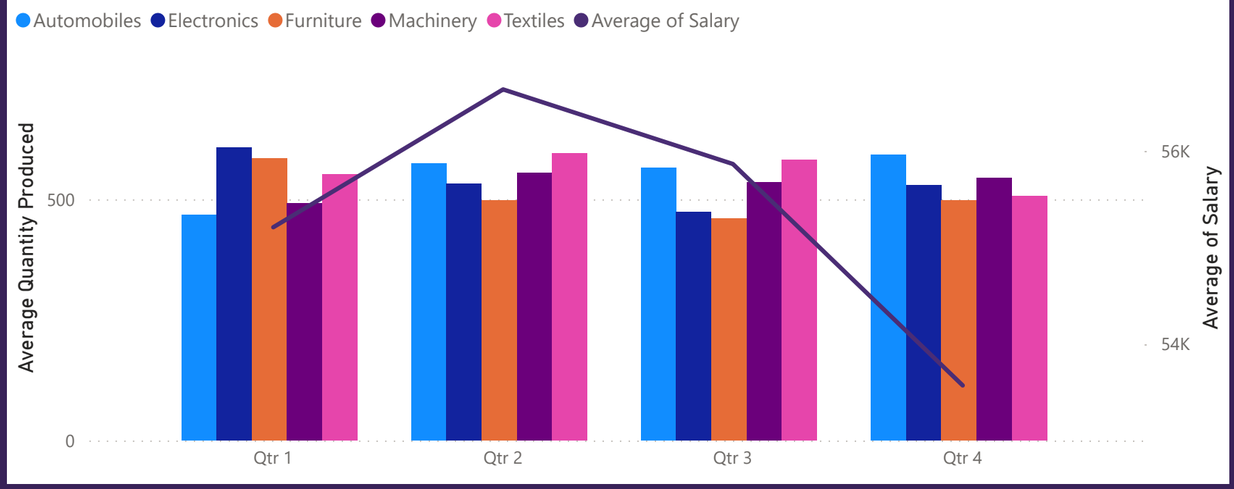
Average Quantity by ProductType



Total Production Quantity over Time



Average Quantity & Salary over the Time



Inferences from the Dashboard visuals:

- Average salary of employee is 55.07K.
- Highest Quantity produced in month of May and lowest in month of December.
- The relationship between production cost and quantity produced is inverse.
- India is producing more textile products, Production cost for electronics is higher in USA
- Percentage of product produced is highest for textiles.
- Average salary of the Production department is maximum followed by Human Resources.
- Maximum quantity of product produced in any quarter is for textile with 33026 quantities being produced.
- USA is the leading producer of goods for the Q2 2025.
- From the graph We can see that the majority of respondents gave their rating 5.
- The majority of employees are from Germany ,followed by India
- The department with the most employees is logistics.