Supply Chain Dataset Analysis – Excel & Power BI Guide

Mapped from Python steps to Excel formulas, PivotTables, Power Query (M), and Power BI DAX/visuals.

Data Overview

Rows: 100 | Columns: 25
Key Columns Detected: Lead time, Production volumes, Manufacturing costs, Defect rates, Transportation modes, Routes, Inspection results.

Metric	Value
Avg Lead time	17.08
Avg Manufacturing costs	47.27
Avg Production volumes	567.84
Avg Defect rates	2.28
Avg Cost per unit	0.11

By Transportation Mode – averages

Transportation Mode	Avg Lead Time	Avg Mfg Cost	Avg Cost/Unit	Count
Air	15.96	42.36	0.1	26
Sea	16.41	40.13	0.08	17
Rail	17.54	49.54	0.11	28
Road	18.03	53.65	0.15	29

Top Routes by Avg Manufacturing Cost (Cheapest first)

Route	Avg Lead Time	Avg Mfg Cost	Avg Cost/Unit	Count
Route C	16.55	41.4	0.08	20
Route B	18.16	46.4	0.12	37
Route A	16.4	50.74	0.12	43

Inspection Result Summary

Inspection Result	Avg Defect	Avg Cost	Count
Fail	2.57	52.23	36
Pass	2.04	46.14	23
Pending	2.15	43.54	41

Correlation between Defect rates and Manufacturing costs: -0.0078 (near zero ⇒ weak linear relationship in this dataset).

Week 1 – Data Preprocessing (Replace Python with Excel & Power BI)

Goal: Prepare clean, structured data.

Excel (Power Query + Formulas):

```
1) Import CSV → Data > Get Data > From Text/CSV.
2) In Power Query:
   • Ensure data types: Transform > Data Type (Lead time/Costs/Volumes as Decimal Number).
   • Standardize 'Inspection results':
       Add Column > Custom Column:
         = Text.Proper(Text.Trim([Inspection results]))
   • Create 'Cost per unit':
       Add Column > Custom Column:
         = if [Production volumes] <> 0 then [Manufacturing costs] / [Production volumes] else null
   • Close & Load to Excel.
3) In Excel sheet (alternative, not Power Query):
   • Standardize inspection text:
       =PROPER(TRIM([@[Inspection results]]))
   • Cost per unit:
       =IFERROR([@[Manufacturing costs]]/[@[Production volumes]],0)
4) Save as a clean table: Format as Table and save workbook.
```

Power BI (Power Query + Model):

```
1) Get data > Text/CSV > supply_chain_data.csv.
2) Power Query steps (Home > Transform Data):
   • Change Type for numeric fields.
   • Add Column (Custom):
       Cost per unit = if [Production volumes] <> 0 then [Manufacturing costs]/[Production volumes] else nu
   • Add Column (Custom):
       Inspection results (clean) = Text.Proper(Text.Trim([Inspection results]))
   • Close & Apply.
3) Create DAX measures:
   Avg Lead Time = AVERAGE('Supply'[Lead time])
   Avg Mfg Cost = AVERAGE('Supply'[Manufacturing costs])
   Avg Cost/Unit = DIVIDE(SUM('Supply'[Manufacturing costs]), SUM('Supply'[Production volumes]))
   Avg Defect Rate = AVERAGE('Supply'[Defect rates])
```

Week 2 – Analysis Questions (Excel & Power BI)

Goal: Answer business questions via PivotTables and DAX.

Excel – PivotTables & Charts

```
Q1) Which transportation mode is fastest vs most expensive?
   Insert > PivotTable (use clean table)
   • Rows: Transportation modes
   • Values: Average of Lead time; Average of Manufacturing costs
   • Sort by Avg Lead time ascending
   • Insert PivotChart: Clustered Columns (or Combo: line for Lead time)

Q2) Do higher defect rates increase manufacturing cost?
   • Insert > Scatter (X: Defect rates, Y: Manufacturing costs)
   • Add Trendline and display R²
   • Compute Pearson correlation:
       =CORREL([Defect rates column], [Manufacturing costs column])

Q3) Which routes are most cost-efficient?
   • PivotTable
   • Rows: Routes
   • Values: Average of Manufacturing costs, Average of Lead time
   • Sort by Avg Manufacturing costs ascending
```

Power BI – Visuals & Measures

```
• Matrix visual (Mode vs Metrics):
   Rows: Transportation modes
   Values: Avg Lead Time (measure), Avg Mfg Cost (measure), Avg Cost/Unit (measure)

• Scatter visual (Defect vs Cost):
   X: Defect rates, Y: Manufacturing costs
   Analytics: Add Trend line (show equation/R² if available in your version)

• Table (Routes by efficiency):
   Columns: Route, Avg Lead Time, Avg Mfg Cost, Avg Cost/Unit
   Sort by Avg Mfg Cost ascending

(Optional) Pearson correlation measure between Defect and Cost:
Corr Defect-Cost =
VAR mx = AVERAGEX(ALLSELECTED('Supply'), 'Supply'[Defect rates])
VAR my = AVERAGEX(ALLSELECTED('Supply'), 'Supply'[Manufacturing costs])
VAR cov = SUMX(ALLSELECTED('Supply'), ('Supply'[Defect rates]-mx)*('Supply'[Manufacturing costs]-my))
VAR sx  = SQRT(SUMX(ALLSELECTED('Supply'), POWER('Supply'[Defect rates]-mx,2)))
VAR sy  = SQRT(SUMX(ALLSELECTED('Supply'), POWER('Supply'[Manufacturing costs]-my,2)))
RETURN DIVIDE(cov, sx*sy)
```

Comparison: Modes by Avg Lead vs Cost

Mode	Avg Lead	Avg Cost	Avg CPU	N
Air	15.96	42.36	0.1	26
Sea	16.41	40.13	0.08	17
Rail	17.54	49.54	0.11	28
Road	18.03	53.65	0.15	29

Comparison: Cheapest Routes (Top 10)

Route	Avg Lead	Avg Cost	Avg CPU	N
Route C	16.55	41.4	0.08	20
Route B	18.16	46.4	0.12	37
Route A	16.4	50.74	0.12	43

Week 3 – Forecasting (No Date column → use Regression/What-If)

Excel – LINEST / SLOPE & INTERCEPT

```
Goal: Forecast Manufacturing costs from Production volumes.

1) Create columns X=Production volumes, Y=Manufacturing costs.
2) Compute regression coefficients:
   • Slope:     =SLOPE(Y_range, X_range)
   • Intercept: =INTERCEPT(Y_range, X_range)
3) Predicted Cost (Y■): in a new column:
   =($Intercept$) + ($Slope$ * [@[Production volumes]])
4) Error metrics (optional):
   • MAE: =AVERAGE(ABS([@[Manufacturing costs]] - [@[Predicted Cost]]))
   • MAPE: =AVERAGE(ABS(([[@[Manufacturing costs]]]-[@[Predicted Cost]])/[@[Manufacturing costs]]))

Lead time from Mode/Route (categorical):
   • Use Pivot averages as baseline targets (lookup by Mode/Route).
   • Or build multiple regression using dummy variables:
     Create 0/1 columns per mode/route and use LINEST with multiple X.
```

Power BI – Measures & What-If Parameter

```
1) Create a What-If parameter for Production volumes (Modeling > New Parameter).
2) Create regression coefficients calculated offline (from Excel) or compute via Quick Measure/measure.
   Example predicted cost measure:
   Pred Cost = [Intercept] + [Slope] * SELECTEDVALUE('Production Volume Parameter'[Production Volume Value]
3) Use a line/column combo visual to show Actual vs Predicted Manufacturing costs.
4) For Lead time: create a table of average Lead time by Mode/Route (Matrix) and use slicers for scenario c
```

Week 4 – Visualization & Storytelling

Excel – Dashboard Steps

Excel Dashboard:
1) KPI Cards: Use cells with formulas (AVERAGE) and Data Bars/Icons (Conditional Formatting).
2) PivotCharts:
   • Combo: Mode vs Avg Lead time (line) & Avg Cost (columns)
   • Bar: Routes sorted by Avg Manufacturing cost
   • Scatter: Defect rate vs Manufacturing cost (add Trendline, show R^2)
3) Slicers: Add slicers for Transportation modes / Routes / Inspection results.

Power BI – Dashboard Steps

Power BI Dashboard:
1) Cards: Avg Lead Time, Avg Mfg Cost, Avg Cost/Unit, Avg Defect Rate (DAX measures from Week 1).
2) Visuals:
   • Matrix: Transportation modes (rows) with Avg Lead, Avg Cost, Avg CPU
   • Bar chart: Routes by Avg Cost (ascending)
   • Scatter: Defect vs Cost, add trendline (Analytics)
   • (Optional) Decomposition Tree: Cost drivers (requires Power BI Desktop)
3) Filters / Slicers: Mode, Route, Inspection result.
4) Layout: Top = KPI cards; Left = Modes matrix; Right = Routes bar; Bottom = Scatter.

Recommendations & Notes

Data-driven Findings (from this dataset):
• Fastest mode: Air (Avg Lead ≈ 15.96); Most expensive mode: Road (Avg Cost ≈ 53.65).
• Cheapest route by cost: Route C (Avg Cost ≈ 41.40); Fastest route: Route A (Avg Lead ≈ 16.40).
• Defect vs cost correlation ≈ -0.0078 → minimal linear relationship; focus on process rather than cost li
• 'Pass' inspections associate with lower average costs and defect levels than 'Fail'.