Introduction

Background:

- This project aims to examine a sample of Sphera ICs (integrated circuits), focusing on potential relationships between **average cost**, **power dissipation**, **process node**, **package type**, and distributors' pricing.

Dataset:

- Contains columns for **Sphera Model Description**, **Pkg Type**, **Node (nm)**, **Power Dissipation (W)**, **Cost** data, and multiple **distributor price** columns.
- Data was compiled from major distributors (Digikey, Arrow, Mouser, Rochester, Avnet) plus manufacturer datasheets for maximum power values, packaging, and node sizes.

Cleaning & Preparation

- **Missing Values**: Filled or dropped for columns such as *Family, Node, Power Dissipation*.
- **Data Type Conversions**: Ensured that numeric columns (e.g., avg node cost) are stored as floats, textual columns (e.g., package type) as strings.
- **Outliers**: Some high-power or high-cost data points remain, which appear legitimate rather than data errors.

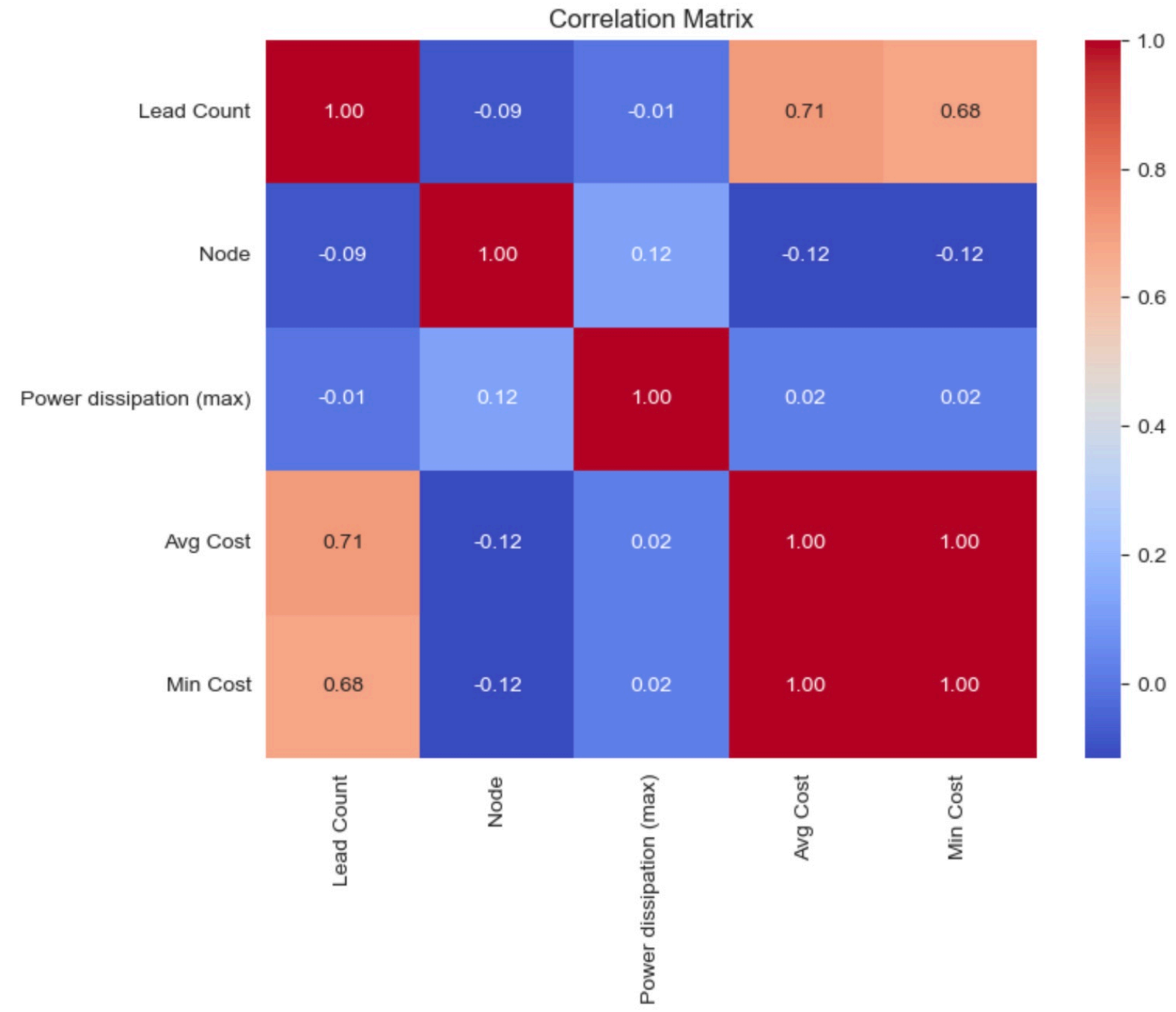
Data Analysis

1. Descriptive Statistics
    - **Power Dissipation**: Mostly under 200W, but a few go beyond 600W, and two over 1000W.
    - **Avg Cost**: Skewed right, with many ICs under $5–$10, but some go over $40 or $100.

	DigiKey (soq)	DigiKey (TR1500 or 1000)	Arrow (soq)	Arrow (1500)	Mouser (soq)	Mouser (1500)	Rochester (1000)	Avnet (1000)
count	35.000000	15.000000	14.000000	12.000000	19.000000	14.000000	22.000000	19.000000
mean	17.362571	13.035143	14.551429	5.477500	17.614737	10.058000	14.850000	13.489263
std	26.107943	18.940770	26.680112	5.042321	27.448645	13.884419	22.911264	18.493795
min	0.380000	0.290000	0.440000	0.390000	0.390000	0.120000	0.090000	0.210000
25%	2.020000	1.290000	1.607500	1.302500	1.780000	0.935500	1.395000	1.540000
50%	6.470000	3.340000	5.060000	3.455000	7.180000	4.150000	5.210000	3.380000
75%	14.400000	13.105000	11.612500	9.167500	16.275000	10.665000	10.697500	13.553000
max	115.580000	59.610000	101.450000	15.510000	114.980000	42.180000	86.390000	58.080000

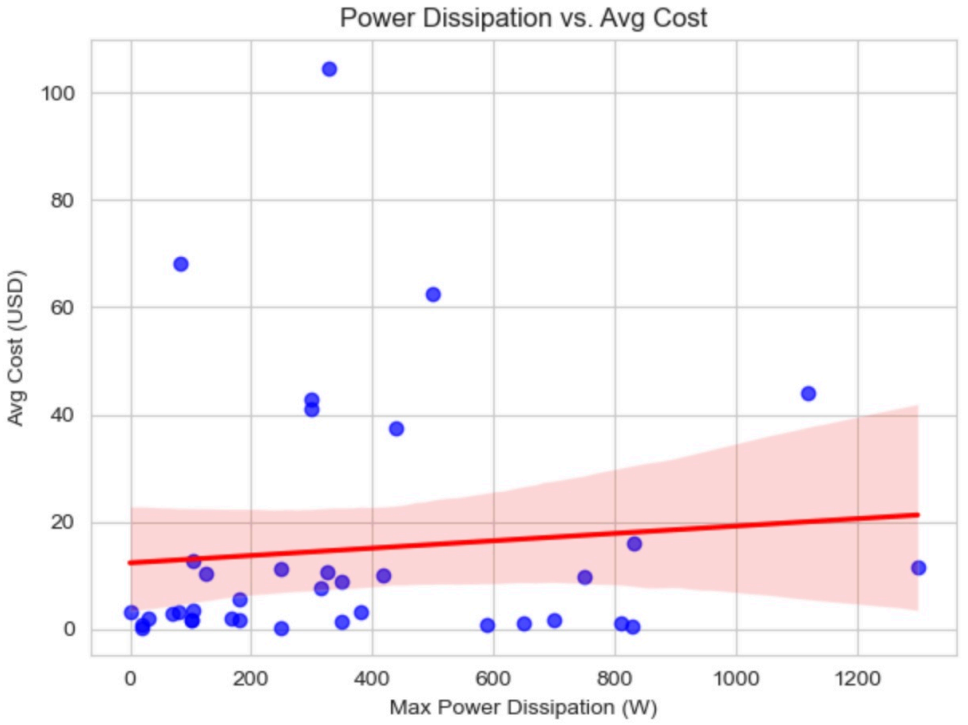
2. Correlation Matrix

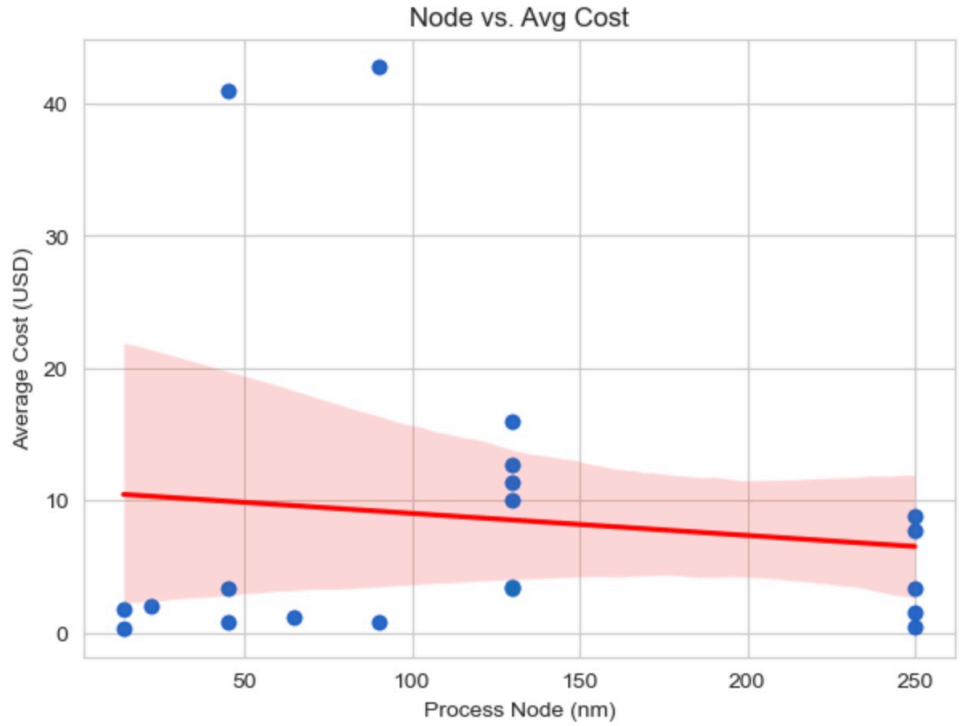
- **Lead Count** strongly correlated with cost (0.68–0.71).
- **Node** negatively correlated with cost (~ -0.12).
- **Power Dissipation** shows negligible correlation (~ 0.02) with cost.



3. Key Plots

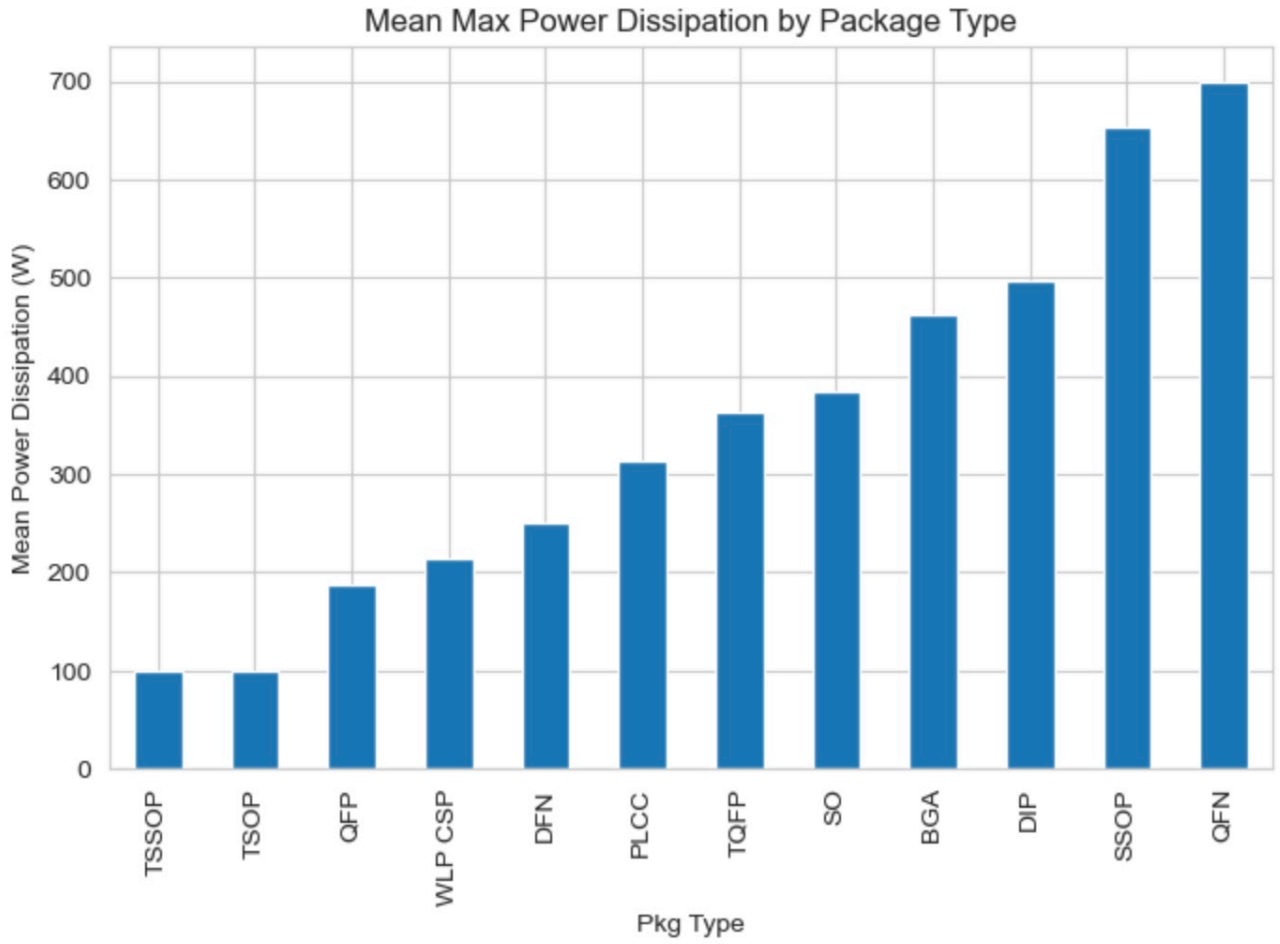
- **Power Dissipation vs. Cost**: Slight upward slope, wide variance.
- **Node vs. Cost**: Slight downward slope, wide variance.
- **Distribution**: Both power and cost heavily skewed with a handful of high-value outliers.





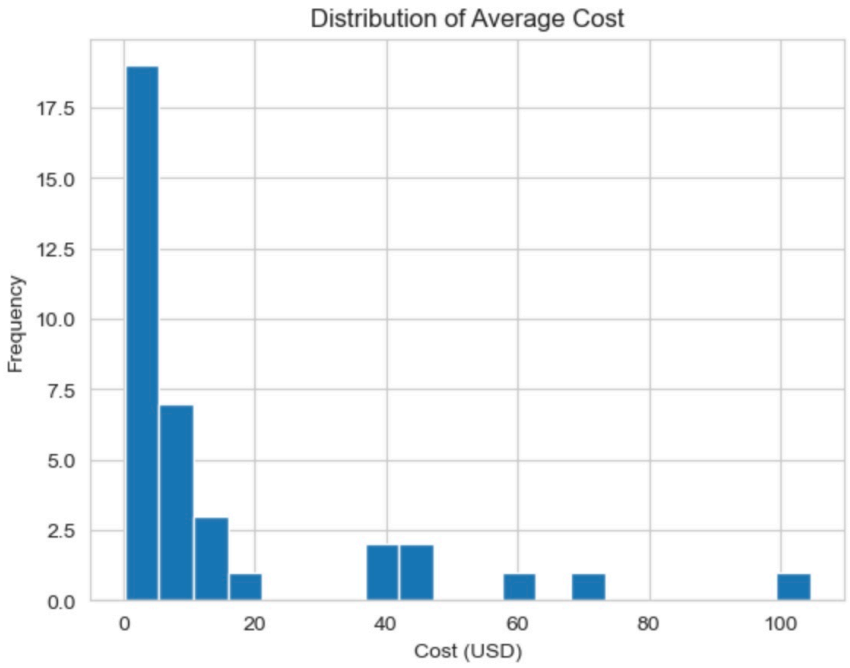
4. Groupby Insights

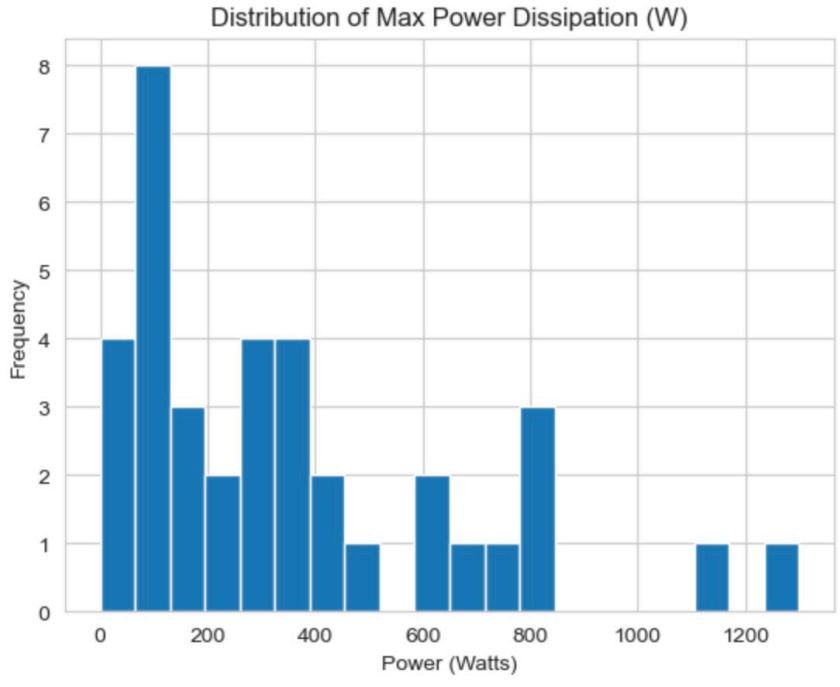
   - **Pkg Type** vs. Mean Max Power: TSSOP and TSOP are on the lower-power end, while packages like QFN, SSOP, DIP, and BGA handle higher power.



5. Distributor Pricing

- Price summary shows different availability or costs across Digikey, Arrow, Mouser, Rochester, and Avnet.
- Some columns have fewer data points, reflecting partial coverage as not every IC was sold by every manufacturer.
- Mean prices range from about $5–$17, with some extremes over $100.





Conclusions & Recommendations

- **Higher Lead Counts → Higher Cost**: This is the strongest correlation found. Designers should note that packages with more leads, often used for complex SoCs, FPGAs, or MCUs, carry higher unit costs.
- **Power Dissipation vs. Cost**: A weak correlation indicates other factors (e.g., performance class, brand, special features) may drive price more than just power rating.
- **Node Technology**: Smaller node processes are mildly associated with higher costs (negative correlation, but not strongly). This aligns somewhat with the intuition that advanced node technology can increase manufacturing costs, though it's not conclusive in this dataset.
- **Distributors**: Pricing can vary significantly. For large-volume orders, analyzing quantity breaks (e.g., TR1500 or 1000) is crucial to find cost savings. Some distributors have outliers, likely reflecting special or hard-to-find parts.