



TENT MARKET ANALYSIS: INSIGHTS FOR PRODUCT & PRICING STRATEGY

JJ Rañeses
8/10/2025

OUTLINE

- Executive Summary
- Introduction
- Methodology
 - Data Review & Cleaning
 - Exploratory Data Analysis (EDA)
 - Comparisons & Segmentation
- Discussion & Findings
 - Exploratory Data Analysis
 - Market Overview
 - Pricing Landscape
 - Product Mix and Within Premium
 - Size, Price & PPSF
 - Other Factors
- Conclusion
- Recommendation
- Appendix



EXECUTIVE SUMMARY

- Dataset covers 721 tent listings across multiple types, sizes, and countries (US & Canada).
- Canadian tents show a consistent price premium, even within the same tent type.
- Product mix, size, and specific features (return policy, height, assembly time) strongly influence price.
- Larger tents cost more in total but can be cheaper per square foot, offering better value.
- Several data quality issues (inconsistent formats, missing values) limit deeper analysis in some dimensions.



INTRODUCTION

Purpose:

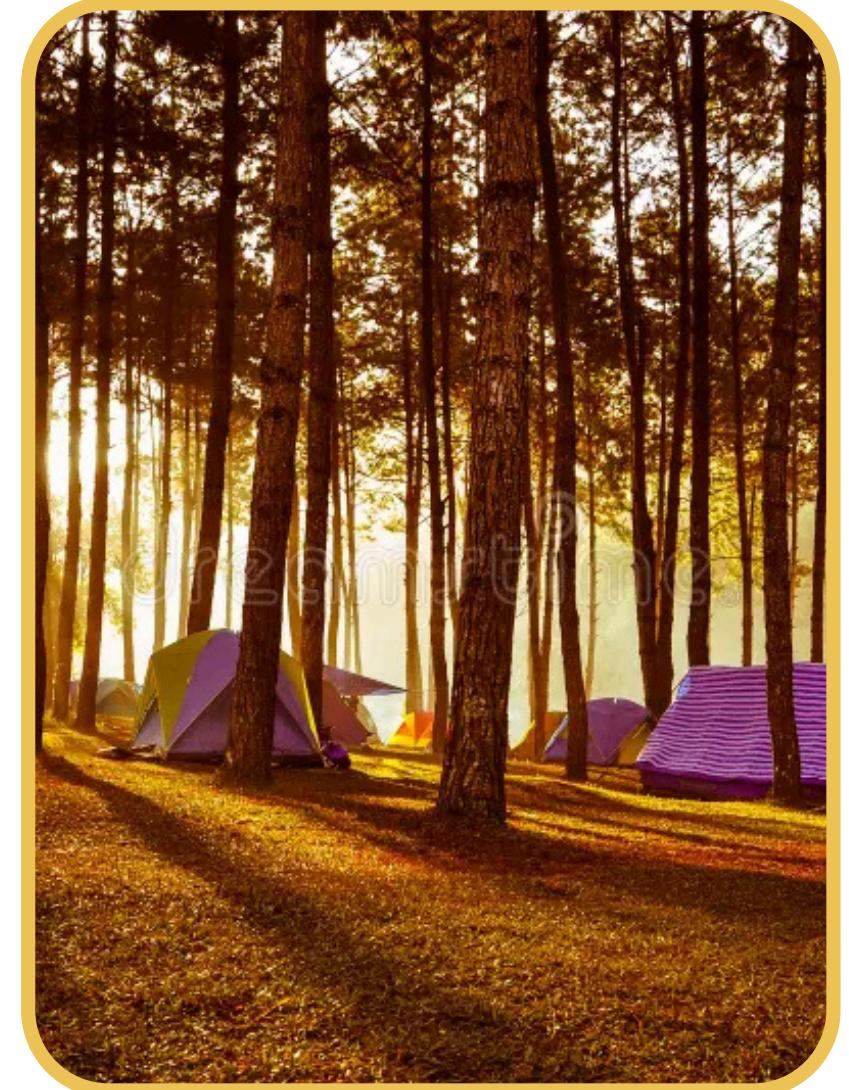
- Understand the tent market structure, price drivers, and cross-country differences.
- Identify trends that can inform pricing strategy, product mix, and market positioning.

Dataset Overview:

- 48 original columns (size, materials, durability, price, etc.).
- 721 total listings, mainly from the US (83%) and Canada (17%).
- Wide variety of tent types – Wall, Teepee, Safari dominate.

Time Spent:

- 8 hrs. for cleaning and analysis and 4 hrs. for presentation



METHODOLOGY



METHODOLOGY

DATA REVIEW & CLEANING

```
# Load the Excel File
from google.colab import drive
drive.mount('/content/drive')
file_path = "/content/drive/MyDrive/Tents_Dataset.xlsx"
tent = pd.ExcelFile(file_path)
df = pd.read_excel(file_path)
df.replace("Not Specified", np.NA)
df.columns = df.columns.str.replace(r"\s+", " ", regex=True).str.strip()
df.head()
```

Product Type	Location	Manufacturer	Model	Website	Link to Product Model	Price	Size (Diameter, ft)	Size (Length, ft)	Size (Width, ft)
0 Teepee	US	Secret Creek	12'	https://coloradoyurt.com/	https://coloradoyurt.com/tipi-inspiration-gall...	733	12	<NA>	<NA>
1 Teepee	US	Secret Creek	16'	https://coloradoyurt.com/	https://coloradoyurt.com/tipi-inspiration-gall...	1070	16	<NA>	<NA>
2 Teepee	US	Secret Creek	18'	https://coloradoyurt.com/	https://coloradoyurt.com/tipi-inspiration-gall...	1407	18	<NA>	<NA>
3 Teepee	US	Secret Creek	20'	https://coloradoyurt.com/	https://coloradoyurt.com/tipi-inspiration-gall...	1685	20	<NA>	<NA>
4 Teepee	US	Secret Creek	24'	https://coloradoyurt.com/	https://coloradoyurt.com/tipi-inspiration-gall...	2247	24	<NA>	<NA>

5 rows × 10 columns

Removed “Not Specified” placeholders → NaN.

```
def undo_excel_date_ranges(series):
    s = series.copy()
    dt = pd.to_datetime(s, errors='coerce')
    mask = dt.notna()
    s = s.astype(str)
    # Cross-platform month/day without leading zeros:
    month = dt.dt.month.astype('Int64').astype(str)
    day = dt.dt.day.astype('Int64').astype(str)
    s.loc[mask] = (month[mask] + "-" + day[mask])
    return s

# Apply to columns that accidentally became dates:
date_broken_cols = ["Room Capacity", "Lead Time (in days)"]
for col in date_broken_cols:
    if col in df.columns:
        df[col] = undo_excel_date_ranges(df[col])

def parse_numeric_or_midpoint(val):
    if pd.isna(val): return np.nan
    text = str(val).strip()
    if text == "": return np.nan

    # Try ranges first: -, en/em dash, 'to', '~, '±', 'approx'
    m = re.search(r"(-|en|em|dash|'to'|'~'|'±'|'approx')\s*(?:-|-|to|-|z|=?|\approx|\.)\s*(\d+(:\.\d+)?)", text, flags=re.I)
    if m:
        a, b = float(m.group(1)), float(m.group(2))
        return (a + b) / 2.0

    # Fallback: first numeric token in the string (handles '8-10 years', '50-60 mph', '$1,299')
    m2 = re.search(r"(\d+(:\.\d+)?)", text)
    return float(m2.group(1)) if m2 else np.nan

range_candidates = [
    "Price", "Size (Diameter, ft)", "Size (Length, ft)", "Size (Width, ft)", "Maximum Height (ft)",
    "Area", "Room Capacity", "Wind Rating", "Snow Load", "Lead Time (in days)",
    "Assembly Time (in hours)", "Estimated Life Span (in years)", "Return (in days)", "Warranty (in years)"
]
```

# Check the affected column		
Room Capacity	Room Capacity_num	
0 1-1	1.0	
58 1-2	1.5	
59 2-3	2.5	
60 3-4	3.5	
61 4-7	5.5	

46 Price_num	556 non-null	float64
49 Size (Diameter, ft)_num	226 non-null	float64
50 Size (Length, ft)_num	426 non-null	float64
51 Size (Width, ft)_num	421 non-null	float64
52 Maximum Height (ft)_num	286 non-null	float64
53 Area_num	87 non-null	float64
54 Room Capacity_num	176 non-null	float64
55 Wind Rating_num	65 non-null	float64
56 Snow Load_num	19 non-null	float64
57 Lead Time (in days)_num	151 non-null	float64
58 Assembly Time (in hours)_num	88 non-null	float64
59 Estimated Life Span (in years)_num	43 non-null	float64
60 Return (in days)_num	45 non-null	float64
61 Warranty (in years)_num	130 non-null	float64
62 warranty_main_years	56 non-null	float64
63 warranty_fabric_years	1 non-null	float64

Fixed broken numeric values (ranges, date conversions) and created cleaned numeric columns (*_num)

```
# Standardize column names by removing trailing spaces, joining broken names, replaces multiple space with single space
df.columns = df.columns.str.replace(r"\s+", " ", regex=True).str.strip()
```

Standardized column names for ease of reference.

```
if "Size (Length, ft)_num" in df.columns and "Size (Width, ft)_num" in df.columns:
    df["estimated_area_ft2"] = (
        pd.to_numeric(df["Size (Length, ft)_num"], errors="coerce")
        * pd.to_numeric(df["Size (Width, ft)_num"], errors="coerce")
    )

# Calculate price per square area (given that price and area exist, and they are not null)
df["price_per_ft2"] = np.nan
ppsa = (
    df["Price_num"].notna() &
    df["estimated_area_ft2"].notna() &
    (df["estimated_area_ft2"] > 0)
)

df.loc[ppsa, "price_per_ft2"] = (
    pd.to_numeric(df.loc[ppsa, "Price_num"], errors="coerce")
    / df.loc[ppsa, "estimated_area_ft2"]
```

count	350.000000	421.000000	550.000000
mean	3563.57723	245.953133	16.916482
std	5101.18160	146.319235	16.830107
min	179.97000	49.000000	4.034392
median	1687.50000	217.562500	10.656250

Added calculated metrics like price_per_ft² and estimated_area_ft².

```
missing = df.isna().sum()
missing = missing[missing > 0].sort_values(ascending=False)

missing_pct = (missing / len(df) * 100).round(1)

missing_summary = pd.DataFrame({
    "missing_count": missing,
    "missing_pct": missing_pct
})
```

warranty_fabric_years	720	99.9
Snow Load_num	702	97.4
Snow Load	698	96.8
Bedroom	679	94.2
Estimated Life Span (in years)_num	678	94.0
After-sales Support	678	94.0
Return (in days)	676	93.8
Return (in days)_num	676	93.8
Kitchen Type	673	93.3

Check for missing values

METHODOLOGY

EXPLORATORY DATA ANALYSIS (EDA)

```

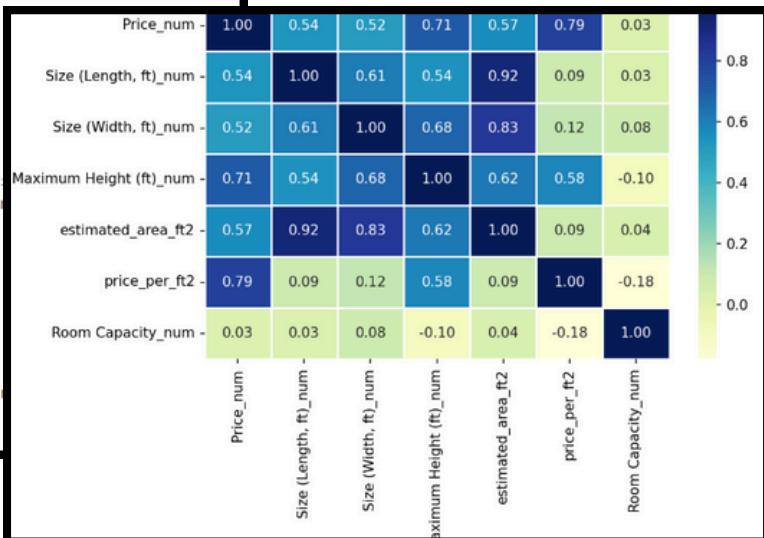
analysis_num_cols = [
    "Price_num",
    "Size (Length, ft)_num",
    "Size (Width, ft)_num",
    "Maximum Height (ft)_num",
    "estimated_area_ft2",
    "price_per_ft2",
    "Room Capacity_num"
]

# Filter only existing columns and drop rows with all NaNs in the analysis
analysis_num_cols = [c for c in analysis_num_cols if c in df.columns]
df = df[analysis_num_cols].dropna(how="all")

# Compute correlation
corr_matrix = corr_df.corr()

# Plot heatmap
plt.figure(figsize=(8,6))
sns.heatmap(corr_matrix, annot=True, cmap="YlGnBu", fmt=".2f", linewidths=1)
plt.title("Correlational Heatmap of Numeric Columns")
plt.show()

```



Finding correlations within numerical columns

```

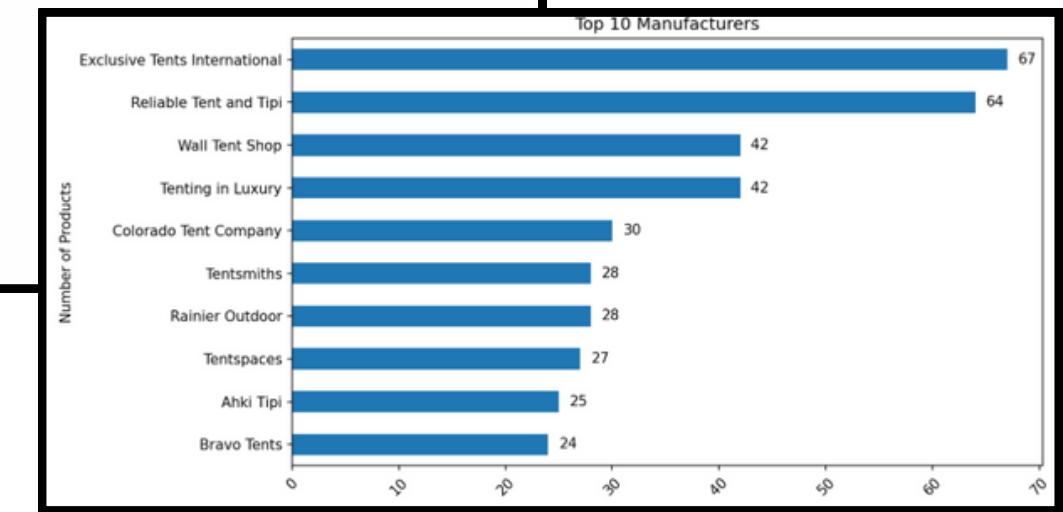
# Determining the top 10 Manufacturers
plt.figure(figsize=(10, 5))
manufacturer = df["Manufacturer"].value_counts().head(10).sort_values(ascending=True)

ax = manufacturer.plot(kind="barh")

for i, v in enumerate(manufacturer):
    ax.text(v + 1, i, str(v), va="center")

plt.title("Top 10 Manufacturers")
plt.ylabel("Number of Products")
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()

```



Delving more on Manufacturers

```

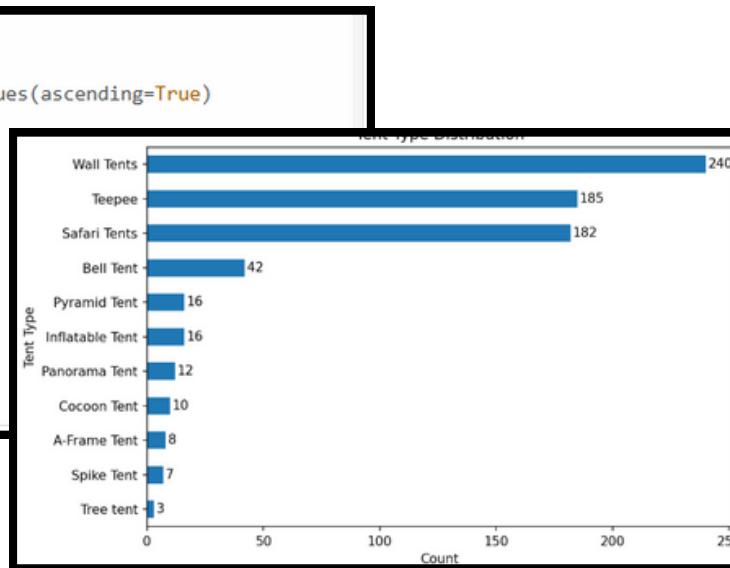
# Determining tent type distribution
plt.figure(figsize=(8,5))
tent = df["Specific Product Type"].value_counts().sort_values(ascending=True)

ax = tent.plot(kind="barh")

for i, v in enumerate(tent):
    ax.text(v + 1, i, str(v), va="center")

plt.title("Tent Type Distribution")
plt.xlabel("Count")
plt.ylabel("Tent Type")
plt.tight_layout()
plt.show()

```



Checking the distribution of tents in the dataset

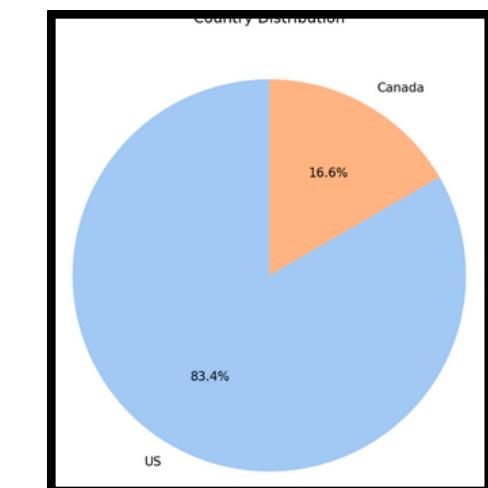
```

country_summary = pd.DataFrame({
    "count": country_counts,
    "percentage": country_pct
})

print("Country Distribution:\n", country_summary)

# Pie chart
plt.figure(figsize=(6, 6))
plt.pie(
    country_counts,
    labels=country_counts.index,
    autopct="%1.1f%%",
    startangle=90,
    colors=sns.color_palette("pastel")[:len(country_counts)])

```



Checking the distribution of countries in the dataset

METHODOLOGY COMPARISON & SEGMENTATIONS

```
# Price-by-Tent-Type in both their total price and price per ft2

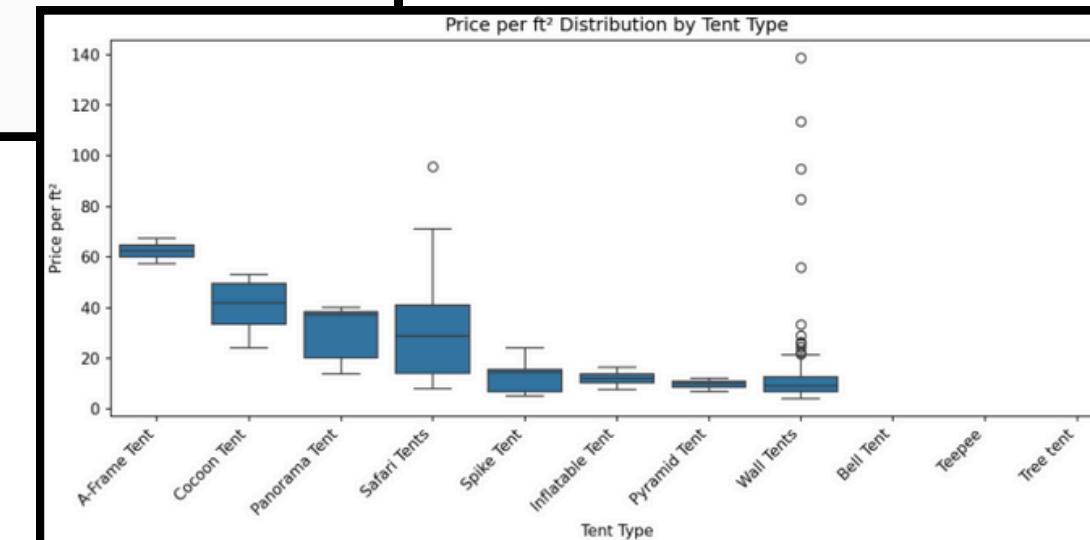
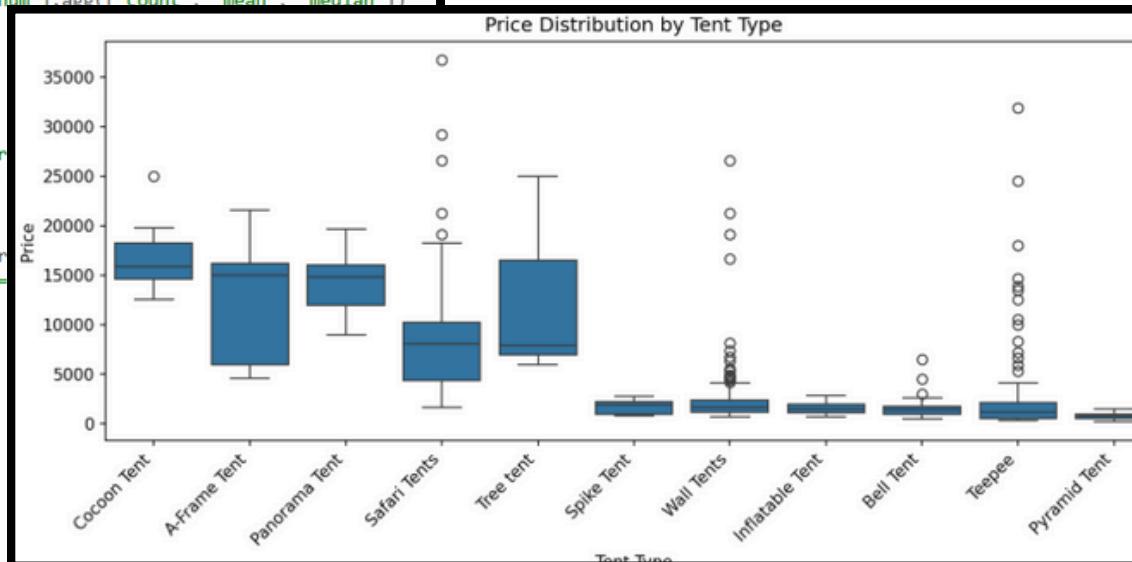
# Price Summary
price_summary = (
    df.groupby("Specific Product Type")["Price_num"].agg(["count", "mean", "median"])
    .round(2).sort_values("median", ascending = False)
)

# Price per ft2 Summary
ppsf_summary = (
    df.groupby("Specific Product Type")["price_per_ft2"].agg(["count", "mean", "median"])
    .round(2).sort_values("median", ascending = False)
)

print("== Price Summary by Tent Type ==\n", price_summary)
print("== Price per ft2 Summary by Tent Type ==\n", ppsf_summary)

# Plot: Price
plt.figure(figsize =(10,5))
sns.boxplot(
    data = df,
    x = "Specific Product Type",
    y = "Price_num",
    order = price_summary.index
)

plt.xticks(rotation = 45, ha = "right")
plt.title("Price Distribution by Tent Type")
plt.ylabel("Price")
plt.xlabel("Tent Type")
plt.tight_layout()
plt.show()
```



Distribution of prices per tent types vs price per ft²

```
# Price by Country

country_summary = (
    df.groupby("Location").agg(
        count = ("Price_num", "median"),
        median_price = ("Price_num", "median"),
        median_ppsf = ("price_per_ft2", "median")
    )
    .reset_index()
)

country_summary = country_summary.sort_values("median_price", ascending = False)

print(f"Price by Country")
country_summary
```

Location	count	median_price	median_ppsf
Canada	1792.0	1792.0	14.057292
US	1682.0	1682.0	10.332026

Median prices between US & Canada

```
# Price & PPSF by Tent Type within Each Country

pivot_price = type_country_summary.pivot(
    index="Specific Product Type", columns="Location", values="median_price"
)
pivot_ppsf = type_country_summary.pivot(
    index="Specific Product Type", columns="Location", values="median_ppsf"
)

print("Median Price by Tent Type & Country")
display(pivot_price)

print("Median Price per ft2 by Tent Type & Country")
display(pivot_ppsf)
```

Location	Canada	US
Specific Product Type		
A-Frame Tent	21576.84	10425.0
Bell Tent	1385.00	1370.0
Cocoon Tent	NaN	15850.0
Inflatable Tent	NaN	14690.0
Panorama Tent	NaN	14800.0
Pyramid Tent	NaN	800.5
Safari Tents	15130.00	6900.0
Spike Tent	NaN	1899.0
Teepee	1749.50	1060.0
Tree tent	25000.00	6925.0
Wall Tents	1440.00	1695.0

Location	Canada	US
Specific Product Type		
A-Frame Tent	67.427625	59.903846
Bell Tent	NaN	NaN
Cocoon Tent	NaN	41.80602
Inflatable Tent	NaN	12.198433
Panorama Tent	NaN	37.215909
Pyramid Tent	NaN	10.069375
Safari Tents	34.972254	28.125
Spike Tent	NaN	14.279365
Teepee	NaN	NaN
Tree tent	NaN	NaN
Wall Tents	11.73	8.82716

Difference between the tent type's price and PPSF in the US and Canada

DISCUSSION & FINDINGS



EXPLORATORY DATA ANALYSIS

WHICH VARIABLES SHOULD WE FOCUS ON TO UNDERSTAND THE TENT MARKET?

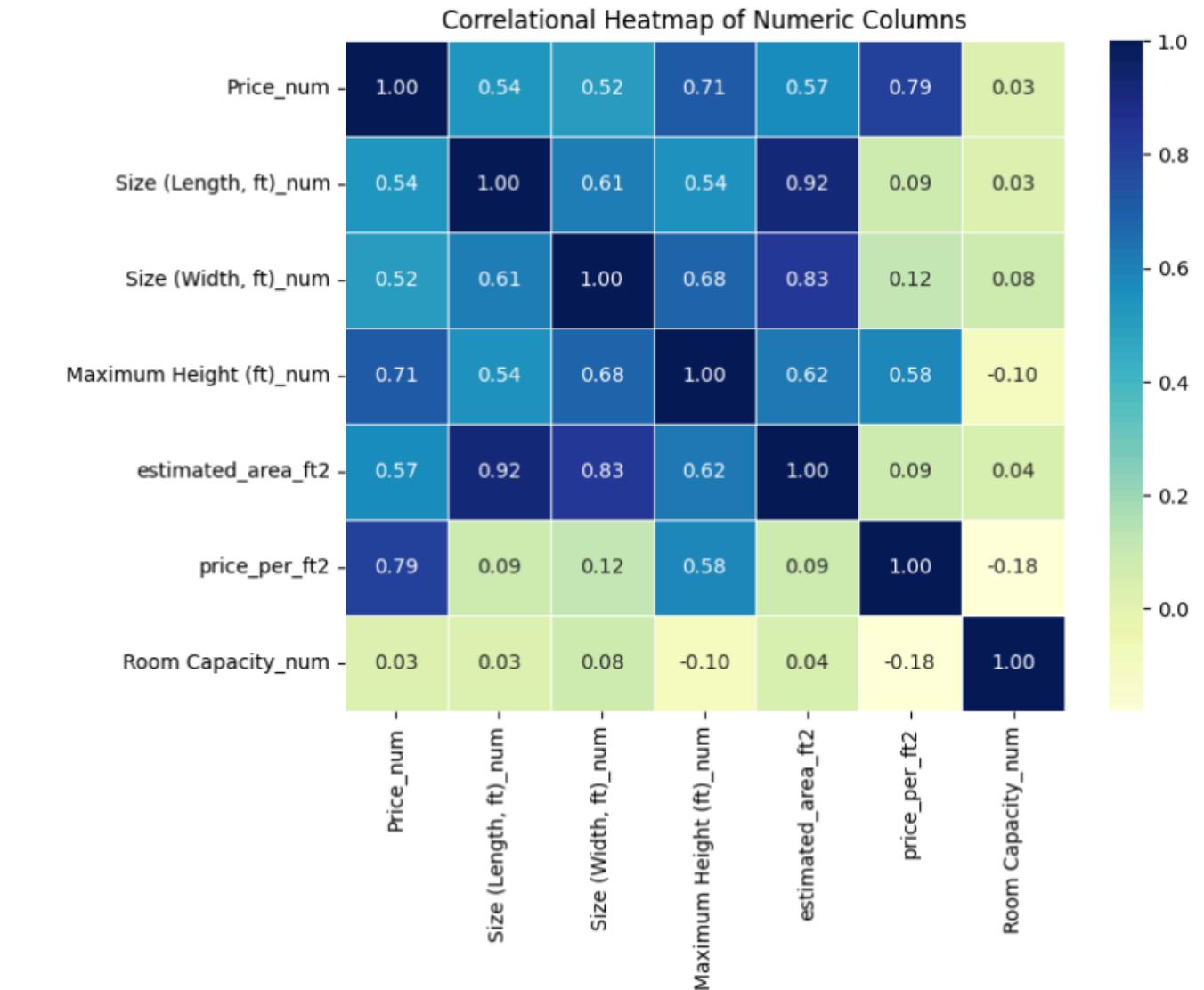
- Price_num is heavily right-skewed: mean (\$3,563) is much higher than median (\$1,688), confirming that median is the more reliable measure for central tendency.
- Area_num has very low coverage (only 87 entries vs 556 for Price), making direct price-per-area analysis incomplete.
- To address missing data:
 - estimated_area_ft2 was created using available Length and Width values.
 - price_per_ft2 (PPSF) was computed from Price_num and estimated_area_ft2 for a consistent value metric across products.
- Assembly Time varies widely (0.24–32 hours), indicating a mix of quick-setup and labor-intensive tents that could influence pricing and buyer decisions.
- Estimated Life Span data is sparse (only 43 entries) but shows most tents last 9–12 years, which could guide warranty positioning.
- Many other numeric attributes (e.g., Wind Rating, Snow Load) also have sparse data, limiting their role in deeper statistical analysis.

	count	mean	std	min	25%	50%	75%	max
Price_num	556.0	3563.577230	5101.181600	179.970000	1039.750000	1687.500000	2948.000000	36750.000000
Size (Diameter, ft)_num	226.0	17.421154	5.425597	8.000000	13.123400	16.40210	20.000000	45.000000
Size (Length, ft)_num	426.0	17.299082	6.567285	7.000000	12.000000	16.000000	20.000000	40.000000
Size (Width, ft)_num	421.0	13.430746	3.663189	6.600000	11.250000	13.000000	16.000000	30.000000
Maximum Height (ft)_num	286.0	9.787581	3.209869	5.249344	8.000000	9.00000	10.000000	32.000000
Area_num	87.0	345.658621	229.733914	56.250000	198.000000	300.000000	443.000000	1060.000000
Room Capacity_num	176.0	3.525568	3.159236	1.000000	1.000000	2.50000	4.625000	13.500000
Wind Rating_num	65.0	54.030769	7.388693	46.000000	46.000000	55.000000	60.000000	75.000000
Snow Load_num	19.0	15.473684	0.512989	15.000000	15.000000	15.00000	16.000000	16.000000
Lead Time (in days)_num	151.0	22.900662	26.507484	1.000000	1.000000	17.50000	35.000000	133.000000
Assembly Time (in hours)_num	80.0	5.166875	8.732585	0.100000	0.237500	0.75000	6.000000	32.000000
Estimated Life Span (in years)_num	43.0	11.104651	4.293625	2.000000	9.000000	10.00000	12.500000	17.500000
Return (in days)_num	45.0	19.000000	6.708204	15.000000	15.000000	15.00000	30.000000	30.000000
Warranty (in years)_num	130.0	7.246154	10.009926	1.000000	2.000000	2.00000	5.000000	30.000000
warranty_main_years	56.0	3.142857	3.288706	1.000000	2.000000	2.00000	3.000000	20.000000
warranty_fabric_years	1.0	7.500000	NaN	7.500000	7.500000	7.50000	7.500000	7.500000
estimated_area_ft2	421.0	245.953133	146.319235	49.000000	144.000000	217.56250	320.000000	1080.000000
price_per_ft2	330.0	16.916482	16.830107	4.034392	7.709772	10.65625	18.206681	138.866667

EXPLORATORY DATA ANALYSIS

WHICH VARIABLES SHOULD WE FOCUS ON TO UNDERSTAND THE TENT MARKET?

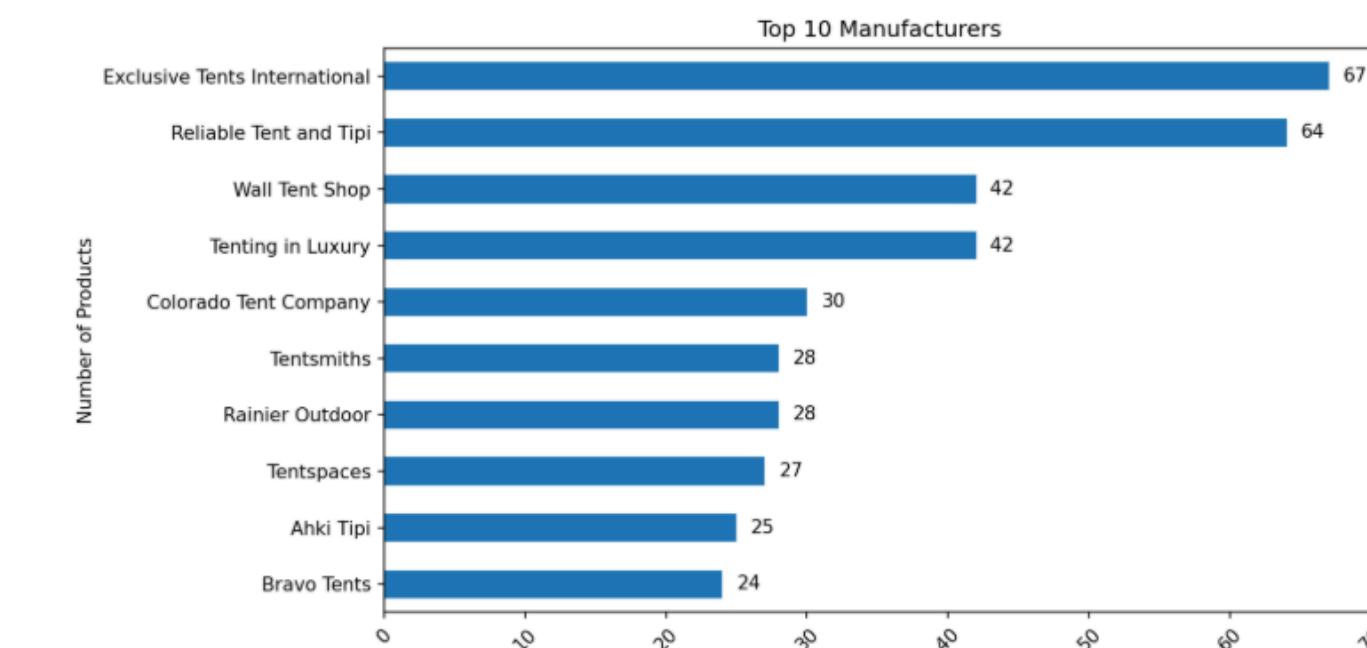
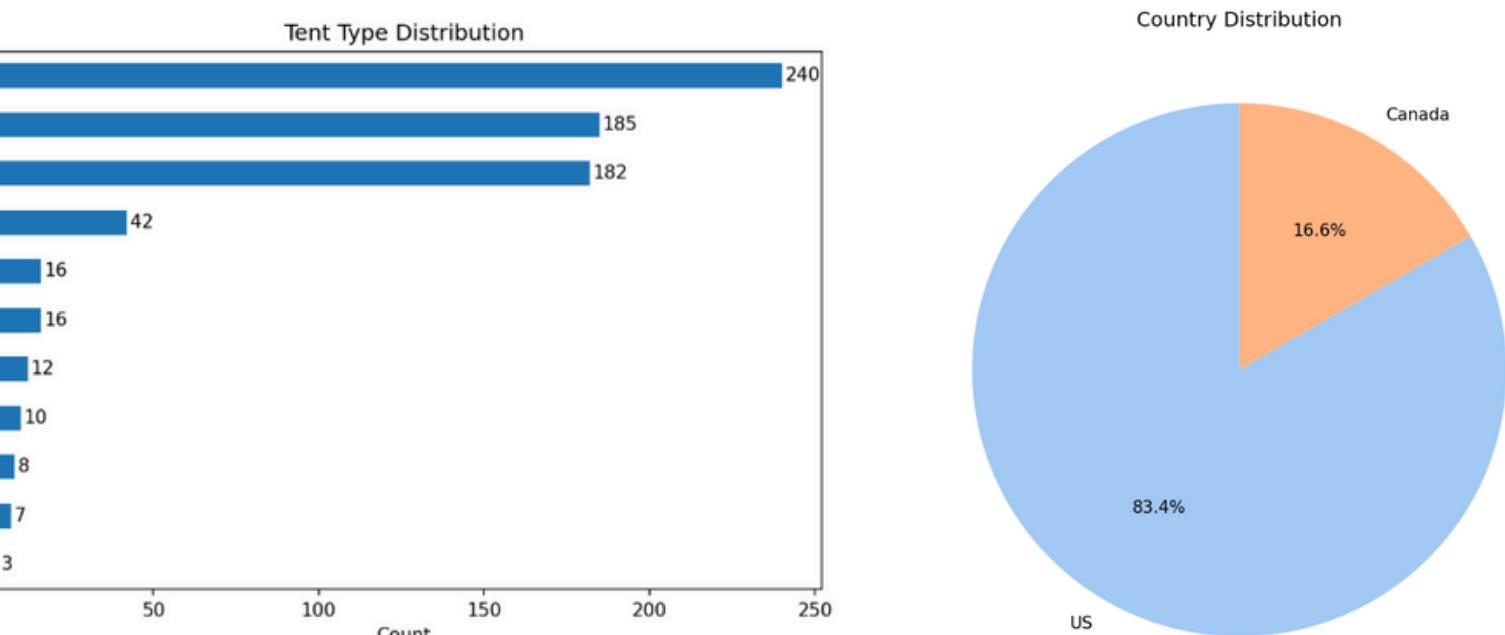
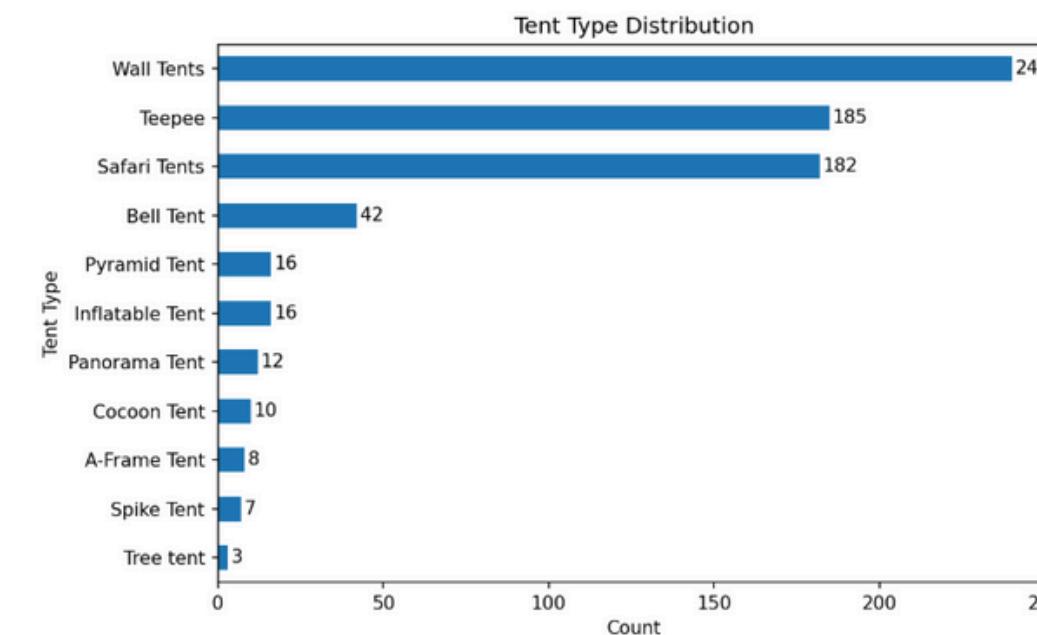
- Price vs Price per ft²: Strong positive correlation (0.79). Higher-priced tents are not only larger but also priced higher per unit area – likely due to premium materials or branding.
- Price vs Maximum Height: High positive correlation (0.71). Taller tents tend to command higher prices, possibly due to added material, engineering costs, or luxury positioning.
- Price per ft² vs Maximum Height: Moderate correlation (0.58). Suggests a “height premium,” where taller tents may cost more even when adjusted for size.
- Room Capacity vs Estimated Area: Very strong correlation (0.92). Bigger tents can host more people, making capacity a reliable size proxy.
- Room Capacity vs Price per ft²: Negative correlation (-0.28). Larger-capacity tents generally have a lower PPSF, reflecting economies of scale.



MARKET OVERVIEW

WHERE DO LISTINGS COME FROM, WHO ARE THE BIGGEST MANUFACTURERS, AND WHICH TENT TYPES DOMINATE?

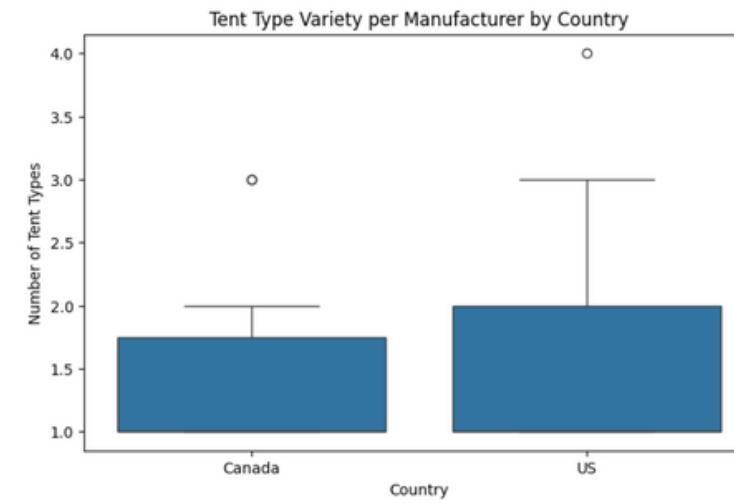
- Origin
 - Over 80% of tents come from the US
 - Canada is the only other major source (~17%)
 - Analysis mainly reflects the US market; price differences may appear when comparing US vs Canada
- Brand Landscape
 - Largest brand sells ~9% of tents in the data
 - Top 5 brands combined account for only ~33%
 - Highly fragmented market; brand alone rarely controls pricing unless it's a premium brand
- Tent Types
 - Wall Tents, Teepees, and Safari Tents make up ~84% of all products
 - Specialty types (Inflatable, Pyramid, Panorama, etc.) are rare, serving niche needs



MARKET OVERVIEW

HOW DO MANUFACTURERS AND PRODUCT VARIETY DIFFER BETWEEN THE US AND CANADA?

	Location	manufacturer_count	avg_tent_types	median_tent_types	max_tent_types
0	Canada	10	1.500000	1.0	3
1	US	43	1.511628	1.0	4



- Canadian manufacturers make up a smaller share of the market and offer fewer tent types on average than US manufacturers.
- Median tent type variety per Canadian manufacturer is lower, pointing to a narrower product mix.
- This supports the idea that Canada's limited tent type availability is partly due to fewer, less diversified manufacturers.

Top 5 Manufacturers per Country by Tent Type Variety

	Location	Manufacturer	tent_types_count
0	Canada	Imago Structures	3
1	Canada	Tenting in Luxury	3
2	Canada	Fort McPherson Tent and Canvas	2
3	Canada	Assiniboine Tipis	1
4	Canada	Biome Canada	1
5	US	Tentspaces	4
6	US	Colorado Tent Company	3
7	US	Montana Canvas	3
8	US	Rainier Outdoor	3
9	US	Secret Creek	3

- US top manufacturers typically offer a broad range of tent types (often 5+), covering multiple market segments.
- Canada's top manufacturers mostly focus on one or two designs, reinforcing the limited variety observed in the Canadian market.

PRICING LANDSCAPE

HOW DO TENT PRICES COMPARE BETWEEN THE US AND CANADA

- Median Price is higher in Canada at 1,792 compared to \$1,682 in the US – a difference of about \$110 ($\approx 6.5\%$ higher). This suggests Canadian listings in this dataset skew toward slightly higher-priced tents.
- Median Price per ft² (PPSF) shows a more pronounced gap:
 - Canada at \$14.06/ft² vs US at \$10.33/ft², a difference of \$3.73/ft² ($\approx 36\%$ higher). This indicates that tents in Canada are either smaller in footprint but higher in quality/specs or simply priced higher per unit of area.
- Both countries have a strong number of listings (Canada: 1,792, US: 1,682), so these comparisons are statistically reliable within this dataset.

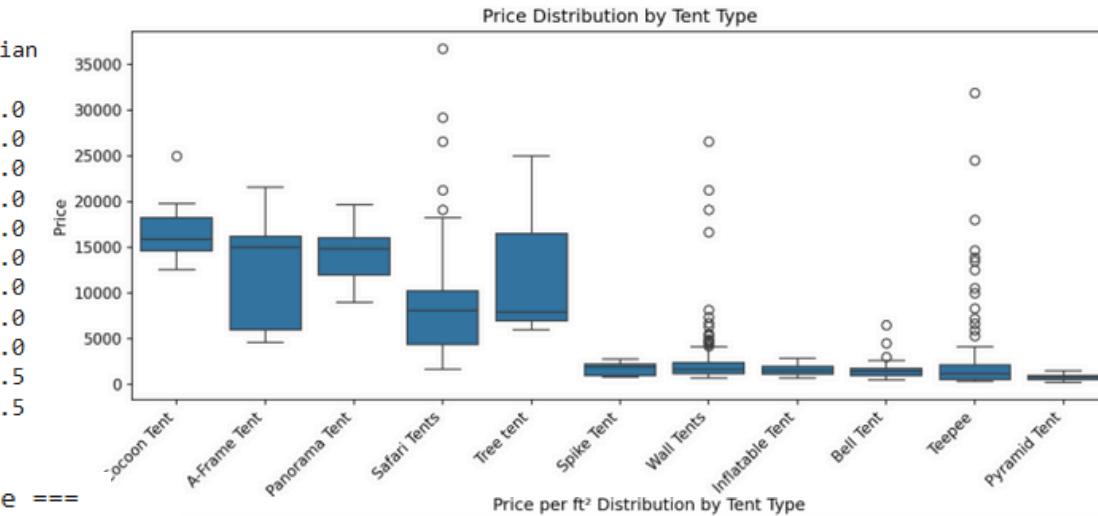
	Location	count	median_price	median_ppsf
0	Canada	1792.0	1792.0	14.057292
1	US	1682.0	1682.0	10.332026

PRICING LANDSCAPE

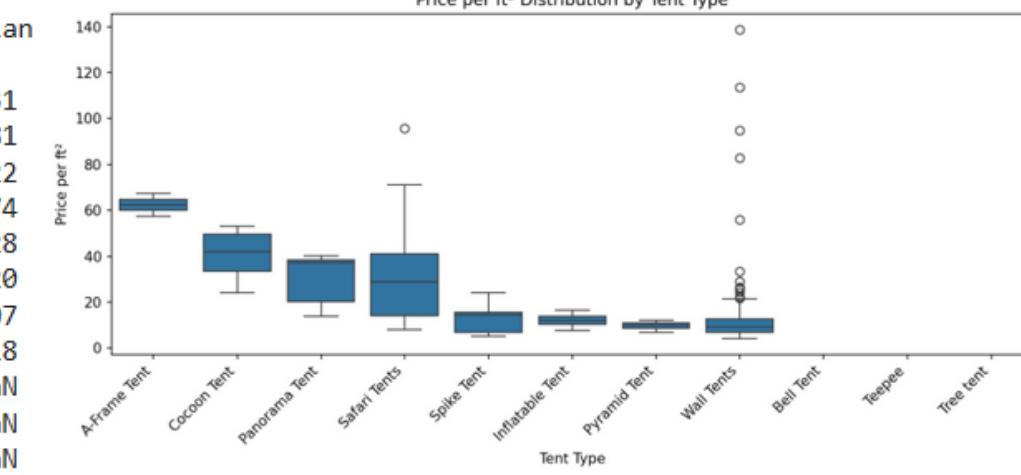
HOW DO PRICES AND VALUE PER SQUARE FOOT VARY ACROSS TENT TYPES?

- High-price niches (Cocoon, A-Frame, Panorama) command strong prices but have very small market share – appealing to specialized buyers, not the mass market.
- Safari Tents lead the mainstream market in price per ft², offering the best balance of volume and high margins.
- Wall Tents dominate in volume at low cost per ft² – ideal for value-focused buyers and volume strategies.
- Mid-tier tents (Teepees, Bell Tents) lack complete size data, limiting pricing analysis – filling this gap could reveal new opportunities.

== Price Summary by Tent Type ==			
Specific Product Type	count	mean	median
Cocoon Tent	10	16755.00	15850.0
A-Frame Tent	5	12645.37	14950.0
Panorama Tent	9	14218.33	14800.0
Safari Tents	64	9170.69	8050.0
Tree tent	3	12950.00	7950.0
Spike Tent	7	1663.86	1899.0
Wall Tents	226	2266.70	1659.0
Inflatable Tent	12	1527.33	1469.0
Bell Tent	42	1587.98	1370.0
Teepee	162	2319.57	1206.5
Pyramid Tent	16	757.42	800.5



== Price per ft ² Summary by Tent Type ==			
Specific Product Type	count	mean	median
A-Frame Tent	3	62.41	62.31
Cocoon Tent	7	40.72	41.81
Panorama Tent	9	29.82	37.22
Safari Tents	60	29.84	28.74
Spike Tent	7	12.64	14.28
Inflatable Tent	11	12.09	12.20
Pyramid Tent	11	9.65	10.07
Wall Tents	222	12.27	9.18
Bell Tent	0	NaN	NaN
Teepee	0	NaN	NaN
Tree tent	0	NaN	NaN



PRICING LANDSCAPE

HOW DO PRICES AND PPSF VARY BY TENT TYPE IN THE US AND CANADA?

- Clear price premium in Canada for several tent types:
 - A-Frame: Canada \$21,576.84 vs US \$10,425 (+107%)
 - Safari: Canada \$15,130 vs US \$6,900 (+119%)
 - Teepee: Canada \$1,749.50 vs US \$1,060 (+65%)
 - Tree Tent: Canada \$25,000 vs US \$6,925 (+261%)
- Similar or slightly lower in Canada:
 - Wall Tent: \$1,440 vs \$1,695
 - Bell Tent: \$1,385 vs \$1,370
- PPSF (Price per ft²) confirms Canadian premium
 - A-Frame: \$67.43 vs \$59.90
 - Safari: \$34.97 vs \$28.13
 - Wall Tent: \$11.73 vs \$8.83
 - All comparable categories show higher PPSF in Canada
- Limitations
 - Missing data for some tent types (Cocoon, Panorama, Pyramid) prevents full cross-country comparison
- Possible Drivers
 - Canada has more listings in high-value categories (A-Frame, Safari, Tree) and fewer in lower-priced types, indicating a sales mix variance effect
 - Even within the same type, Canada's PPSF is higher, suggesting regional pricing differences beyond product mix.

	Location	Canada	US
Specific Product Type			
A-Frame Tent	21576.84	10425.0	
Bell Tent	1385.00	1370.0	
Cocoon Tent	Nan	15850.0	
Inflatable Tent	Nan	1469.0	
Panorama Tent	Nan	14800.0	
Pyramid Tent	Nan	800.5	
Safari Tents	15130.00	6900.0	
Spike Tent	Nan	1899.0	
Teepee	1749.50	1060.0	
Tree tent	25000.00	6925.0	
Wall Tents	1440.00	1695.0	

	Location	Canada	US
Specific Product Type			
A-Frame Tent	67.427625	59.903846	
Bell Tent	Nan	Nan	
Cocoon Tent	Nan	41.806020	
Inflatable Tent	Nan	12.198433	
Panorama Tent	Nan	37.215909	
Pyramid Tent	Nan	10.069375	
Safari Tents	34.972254	28.125000	
Spike Tent	Nan	14.279365	
Teepee	Nan	Nan	
Tree tent	Nan	Nan	
Wall Tents	11.730000	8.827160	

PRODUCT MIX AND WITHIN CATEGORY PREMIUM

WHAT DRIVES CANADA'S PRICE PREMIUM — PRODUCT MIX OR WITHIN-CATEGORY DIFFERENCES

- Over-represented in Canada
 - Wall Tents make up 47.4% of Canadian listings vs 40.6% overall (+6.8 pp)
 - Safari Tents make up 16.7% of Canadian listings vs 11.5% overall (+5.2 pp)
 - Teepees and Tree Tents have similar or slightly higher shares compared to overall
- Under-represented in Canada
 - No listings for Cocoon, Panorama, Pyramid, Inflatable, or Spike Tents (all higher-priced categories)
 - Bell Tents have a smaller share in Canada (5.1%) compared to overall (7.6%)
- Impact on Prices
 - Negative mix effect: Canada's actual median price is \$1,792, but if its product mix matched the overall dataset it would be \$3,270 (-82% mix effect)
 - US also has a negative mix effect (-56%), but less severe than Canada's
 - This product mix skews toward cheaper tent types, which would normally lower Canada's overall median price
- Offset by Within-Category Premium
 - Despite the cheaper mix, Canada still has consistently higher within-category prices
 - This suggests that regional pricing differences, not just product mix, drive the Canadian premium

	Location		actual_price	adjusted_price	mix_effect_diff	mix_effect_pct
0	Canada	1792.0	3270.203597	-1478.203597	-82.489040	
1	US	1682.0	2629.962230	-947.962230	-56.359229	
	Specific Product Type	count_canada	count_overall	pct_canada	pct_overall	diff_pct
10	Wall Tents	37.0	226	47.435897	40.647482	6.788415
6	Safari Tents	13.0	64	16.666667	11.510791	5.155875
9	Tree tent	1.0	3	1.282051	0.539568	0.742483
0	A-Frame Tent	1.0	5	1.282051	0.899281	0.382771
8	Teepee	22.0	162	28.205128	29.136691	-0.931562
7	Spike Tent	0.0	7	0.000000	1.258993	-1.258993
4	Panorama Tent	0.0	9	0.000000	1.618705	-1.618705
2	Cocoon Tent	0.0	10	0.000000	1.798561	-1.798561
3	Inflatable Tent	0.0	12	0.000000	2.158273	-2.158273
1	Bell Tent	4.0	42	5.128205	7.553957	-2.425752
5	Pyramid Tent	0.0	16	0.000000	2.877698	-2.877698

PRODUCT MIX AND WITHIN CATEGORY PREMIUM

WHAT COULD BE DRIVING CANADA'S HIGHER PRICES WITHIN THE SAME TENT TYPE?

Fewer manufacturers in Canada

- Canada has a smaller pool of manufacturers compared to the US.
- Several high-volume US brands are absent from Canadian listings.

Reduced variety per manufacturer

- Canadian manufacturers tend to offer fewer tent types on average.
- Fewer budget-friendly options may push buyers toward higher-priced models within the same category.

Impact on prices

- While the dataset does not allow a direct test, fewer manufacturers and reduced variety may limit competition, helping sustain higher PPSF within the same tent type.

Top 5 Manufacturers per Country by Tent Type Variety

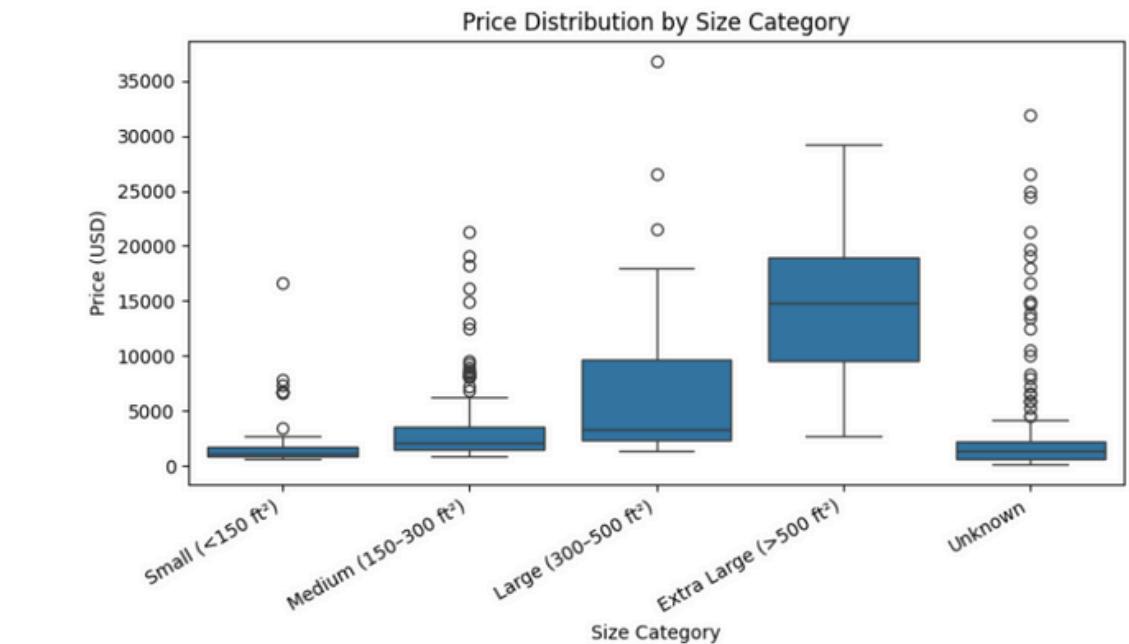
	Location	Manufacturer	tent_types_count
0	Canada	Imago Structures	3
1	Canada	Tenting in Luxury	3
2	Canada	Fort McPherson Tent and Canvas	2
3	Canada	Assiniboine Tipis	1
4	Canada	Biome Canada	1
5	US	Tentspaces	4
6	US	Colorado Tent Company	3
7	US	Montana Canvas	3
8	US	Rainier Outdoor	3
9	US	Secret Creek	3

SIZE, PRICE & PPSF

HOW DOES TENT SIZE AFFECT THE PRICE?

- Tent size is a strong but non-linear driver of price. Median prices climb from \$1,110 for Small tents to \$14,800 for Extra Large tents.
- Upgrading from Small to Medium or Medium to Large comes with only moderate price increases, suggesting mid-range buyers can get more space without a steep cost jump.
- The biggest leap occurs from Large to Extra Large, where prices surge over 4x, reflecting a niche, premium segment aimed at specialized buyers.
- The Unknown size category (median \$1,315) likely contains mostly smaller tents with a few mid-sized ones, explaining its modest price.

Size_Category	count	median	mean
Extra Large (>500 ft ²)	11	14800.0	15052.000000
Large (300–500 ft ²)	78	3273.0	6655.738205
Medium (150–300 ft ²)	133	2130.0	3586.362632
Unknown	226	1315.4	2810.383451
Small (<150 ft ²)	108	1110.5	1708.300648

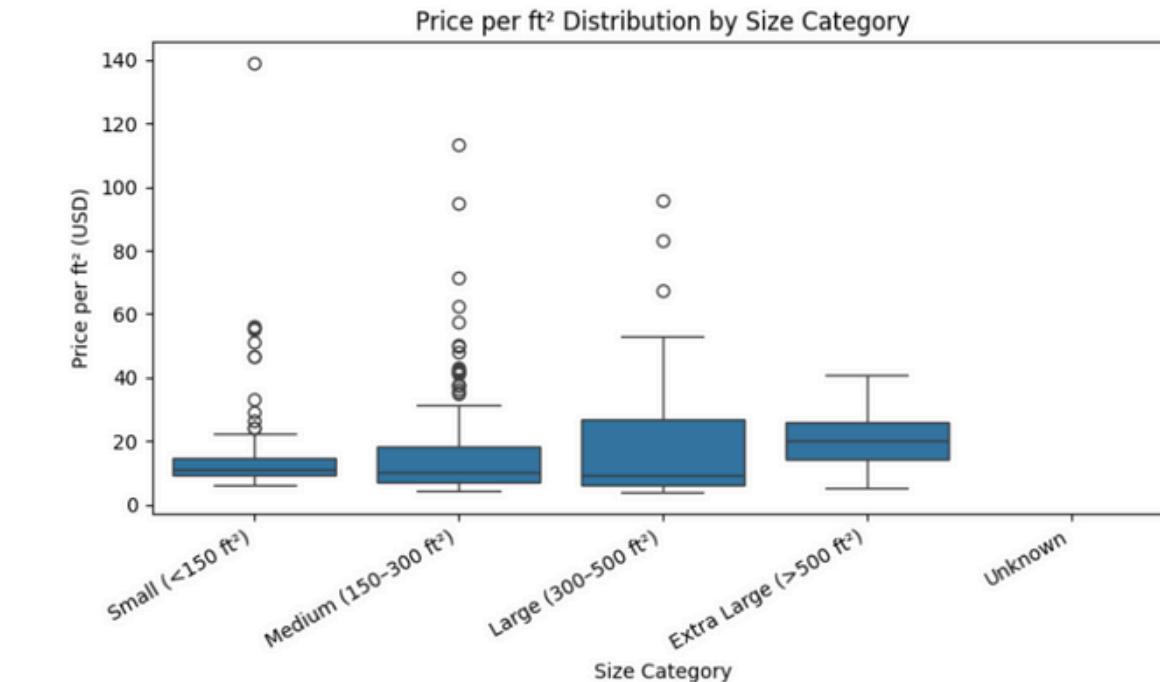


SIZE, PRICE & PPSF

HOW DOES TENT SIZE AFFECT THE PPSF?

- Extra Large tents have the highest median price per ft² (\$20.29/ft²). They are the most expensive both in total price and on a per-unit basis, likely due to premium materials, added engineering, or niche use cases.
- Small, Medium, and Large cluster around \$9.5–\$11.1 per ft², showing little variation in cost efficiency across mid-size categories.
- Large (300–500 ft²) is the cheapest per ft² (\$9.51/ft²), offering the best relative value despite higher total prices.
- Unknown size tents cannot be evaluated per ft²; exclude them from size-based comparisons to avoid distorting averages

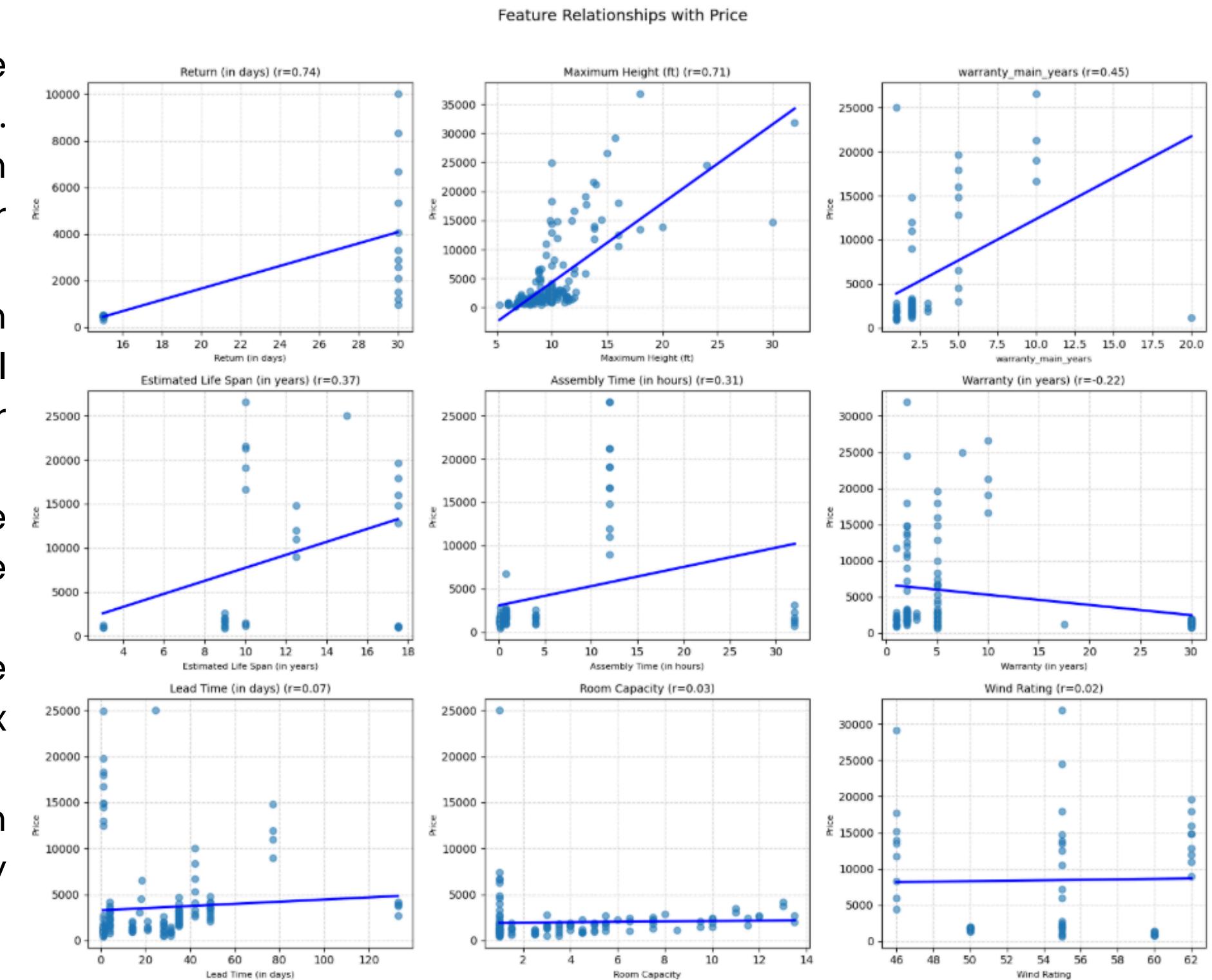
Size_Category	count	median_ppsf	mean_ppsf	median_price
Small (<150 ft ²)	108	11.105525	15.557412	1110.5
Medium (150–300 ft ²)	133	10.029762	16.899738	2130.0
Large (300–500 ft ²)	78	9.516125	18.145393	3273.0
Extra Large (>500 ft ²)	11	20.288248	21.748419	14800.0
Unknown	0	NaN	NaN	1315.4



OTHER FACTORS

WHAT OTHER FACTORS ARE CORRELATED WITH TENT PRICE?

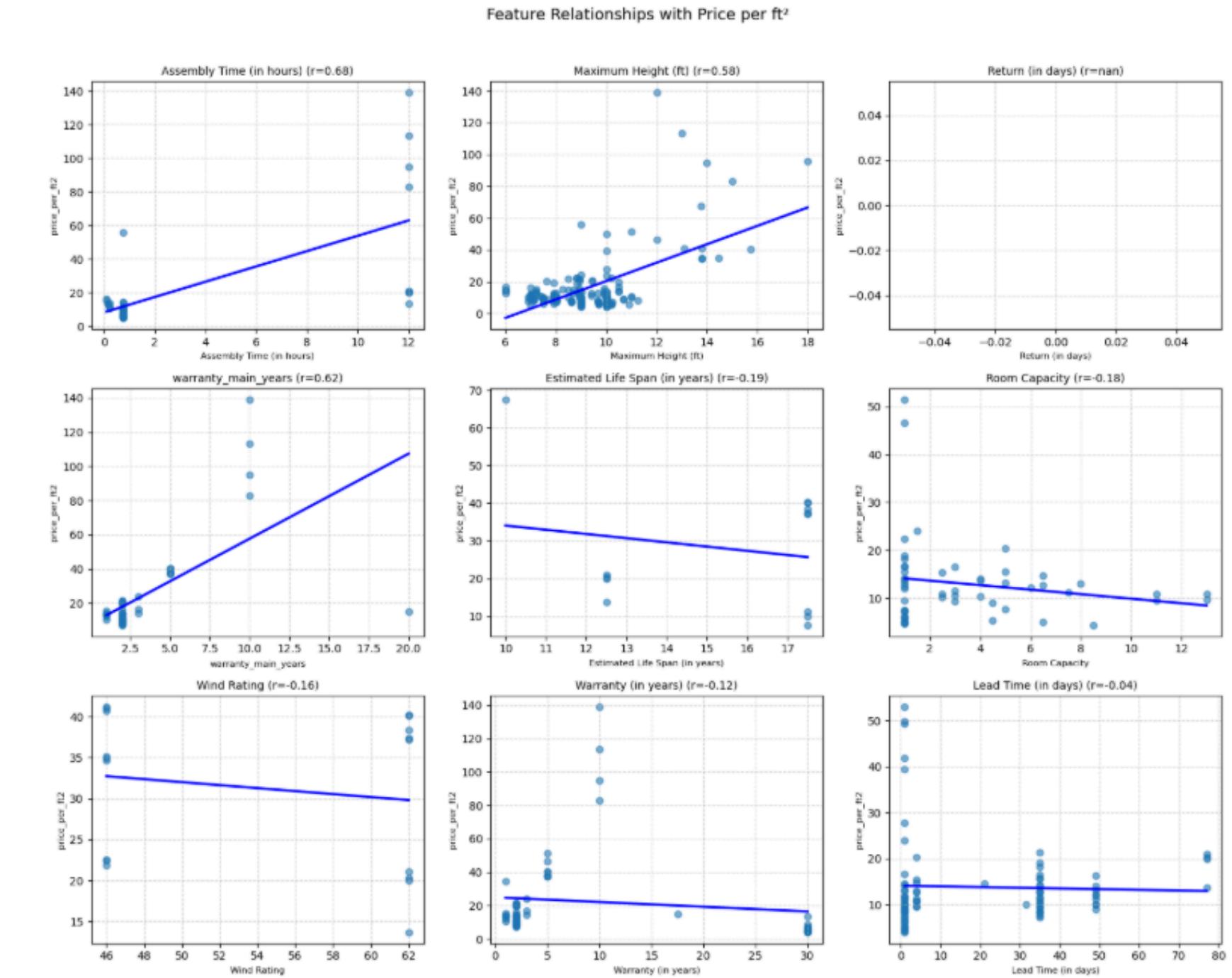
- Return Policy Length – Strong positive correlation (Pearson 0.743, Spearman 0.834). Higher-priced tents often have longer return periods, possibly signaling quality or seller confidence.
- Maximum Height – High positive correlation (0.714). Taller tents command higher total prices, reflecting added material/engineering or luxury positioning.
- Estimated Life Span – Modest positive correlation (0.369). Durability plays some role but likely limited by missing data.
- Assembly Time – Weak-to-moderate positive correlation (0.314). May indicate more complex or niche designs priced higher.
- Little/No Correlation – Lead Time (0.074), Room Capacity (0.030), Wind Rating (0.018), Warranty length (slight negative at -0.225).



OTHER FACTORS

WHAT OTHER FACTORS ARE CORRELATED WITH TENT PPSF?

- Assembly Time – Strong positive correlation (0.679). Longer setup times may be linked to niche or complex tents priced higher per unit area.
- Maximum Height – Moderate positive correlation (0.583). Suggests a “height premium” even after adjusting for size.
- Return Policy Length – Weak positive correlation (0.287). More generous returns add a small premium per square foot.
- Others – Most other features have negligible correlation with PPSF.



CONCLUSION

- Data quality required heavy preparation – Ranges, placeholder values, and date errors were standardized; numeric columns were engineered (e.g., *_num) to enable analysis. Missing data limits certain metrics (e.g., lifespan).
- Market is concentrated in a few tent types – 84% of listings are Wall Tents, Teepees, or Safari Tents; no manufacturer exceeds 9% market share.
- Canada's price gap is driven by within-category premiums – Median price is 6.5% higher, PPSF is 36% higher. Product mix alone would normally lower Canada's prices, but higher pricing for the same tent types offsets this and creates the gap.



CONCLUSION

- Market structure likely contributes – Canada has fewer manufacturers and lower tent-type variety per manufacturer, which may reduce competition.
- Feature-driven premiums are evident – Return Policy Length (Pearson 0.743, Spearman 0.834), Maximum Height (0.714, 0.583), and Assembly Time (0.314, 0.679) are the strongest drivers of price and PPSF.
- Size patterns reinforce value trends – Larger tents cost more in absolute terms but can offer better PPSF value; extra-large tents carry both high total prices and high PPSF.



RECOMMENDATIONS

- Expand manufacturer base in Canada
 - Identify high-volume US manufacturers absent from the Canadian market and explore partnership or distribution opportunities.
 - Goal: Increase competition and variety to reduce within-category premiums.
- Diversify tent-type offerings
 - Add more representation in premium categories (e.g., Cocoon, Panorama) to balance product mix.
 - Target under-represented high-value segments where competitors are scarce.
- Leverage feature-driven pricing
 - Highlight maximum height, assembly complexity, and return policy length in marketing for premium products.
 - For value segments, explore cost efficiencies in these features to offer competitive PPSF.



RECOMMENDATIONS

- Refine pricing strategy for Canada
 - Review pricing for overlapping SKUs with the US, especially in high-volume categories, to assess if the within-category premium aligns with market positioning goals.
- Fill critical data gaps for stronger future analysis
 - Prioritize collecting complete and standardized data for currently sparse attributes such as lifespan, durability specs (snow load, wind rating), and exact tent area.
 - Example: If area is missing, PPSF cannot be calculated for those listings – limiting value and size-based pricing insights.
- Investigate drivers of Canada's within-category premium
 - Match exact models across countries to isolate pricing differences.
 - Capture missing attributes (brand, material quality, shipping cost, retailer type).
 - Measure competition levels per tent type.
 - Add external market factors (exchange rates, tariffs, freight costs) to identify cost-driven effects.
- Improve data collection and standardization at the source
 - Use consistent formats for numeric attributes (avoid ranges like "8–10 years" or placeholders like "Not Specified").
 - Apply automated validation to catch errors like accidental date conversions in numeric fields.



APPENDIX

DummyNRiver/
Tent-Analysis



1 Contributor 0 Issues 0 Stars 0 Forks

Tent-Analysis/JJ_Walden_Tent_Analysis.ipynb at main · DummyNRiver/Tent-Analysis

Contribute to DummyNRiver/Tent-Analysis development by creating an account on GitHub.

[GitHub](#)

https://github.com/DummyNRiver/Tent-Analysis/blob/main/JJ_Walden_Tent_Analysis.ipynb





THANK YOU