
Recommendation for MLB Expansion City

Data-Driven Analysis and Prediction

Ebenezer Olubayode



Problem

Goal: Identifying most profitable and stable intercontinental city for a new MLB franchise via data-driven insights

Success Metrics Could Include:

- Population size and growth potential
- Economic Factors e.g. Median Household Income
- Sports Market Saturation (Other Pro League)
- Existing Baseball Ecosystem (Minor League Teams)

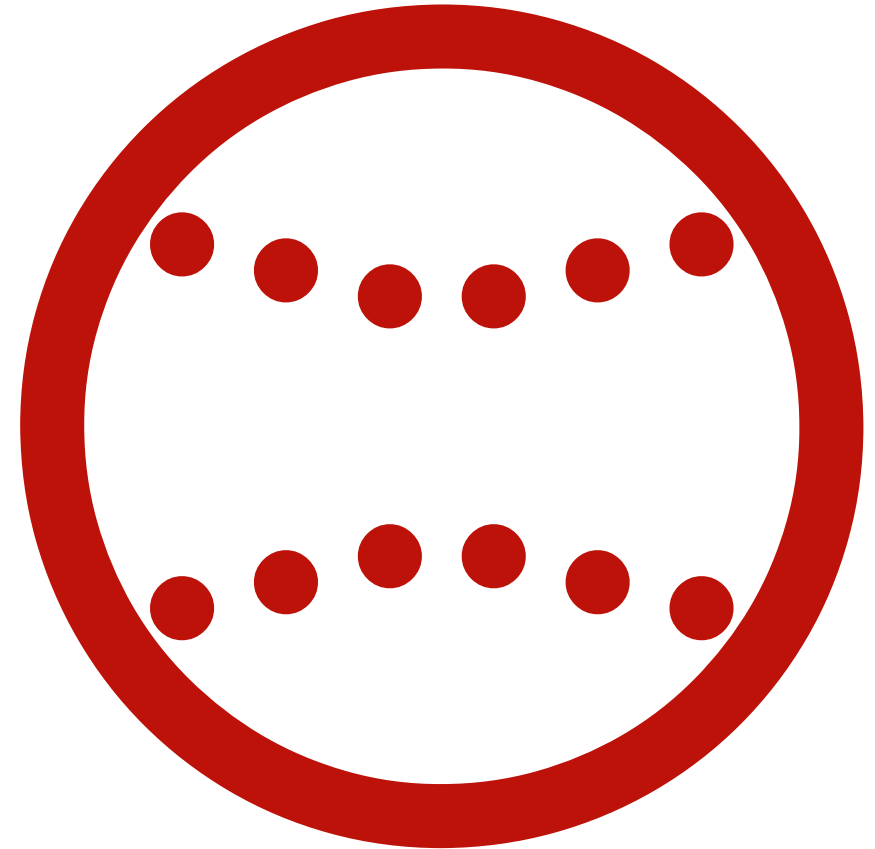


DATA COLLECTION & SOURCES

Data was collected for both **cities with existing 30 MLB teams** and **385 cities with and without MLB teams** to identify patterns and evaluate potential markets for expansion, covering metrics such as:

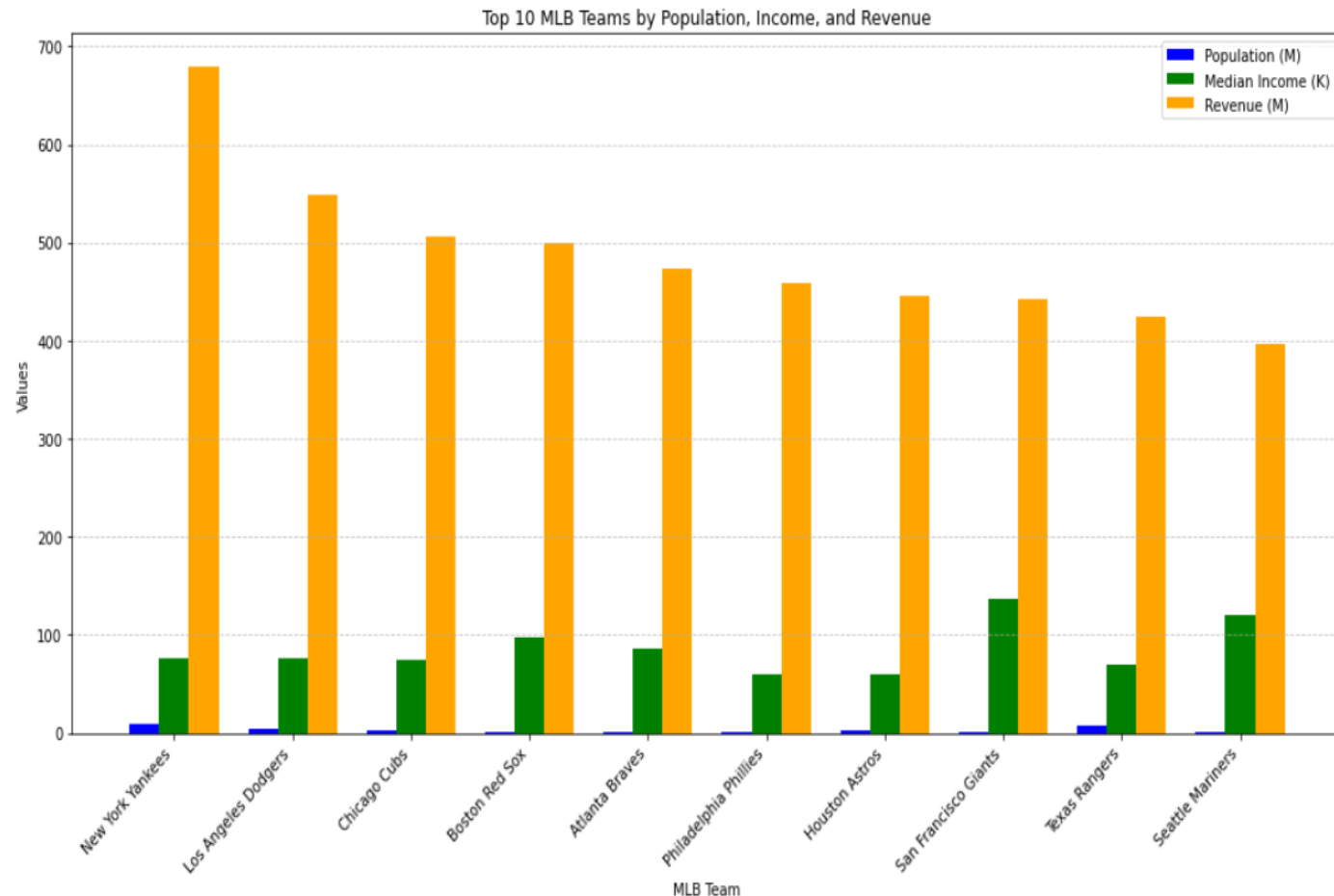
- Population size (city and metro).
- Median household income.
- Total revenue of MLB teams.
- Minor league and other sports teams (NFL, NBA, NHL, MLS)
- Presence of minor league teams (AAA, AA, A).

Sources; Forbes, US Census Data and others



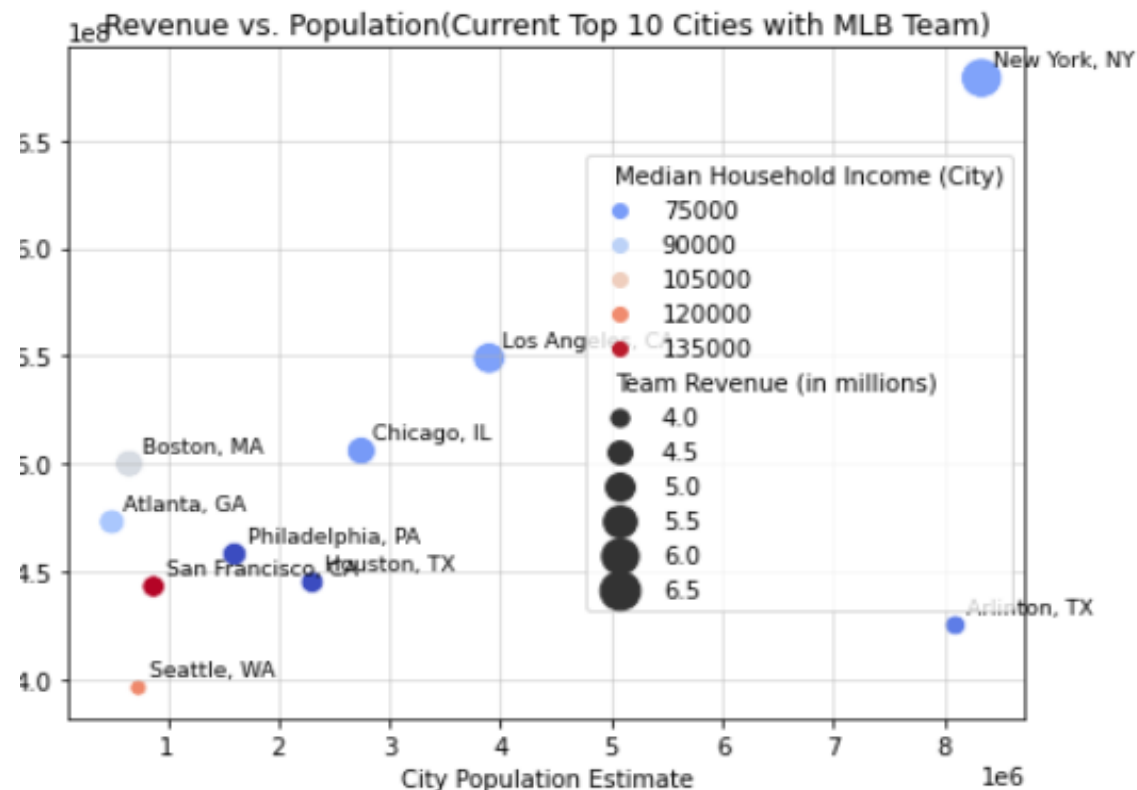
Current Market Size of MLB Teams

- Teams like New York Yankees and Los Angeles Dodgers generate significantly higher revenues, indicating dominant markets.
- Lower revenue team may struggle in smaller or less lucrative markets.
- Revenue is likely influenced by median income and fan engagement and other economic factors



Current Market Size of MLB Teams

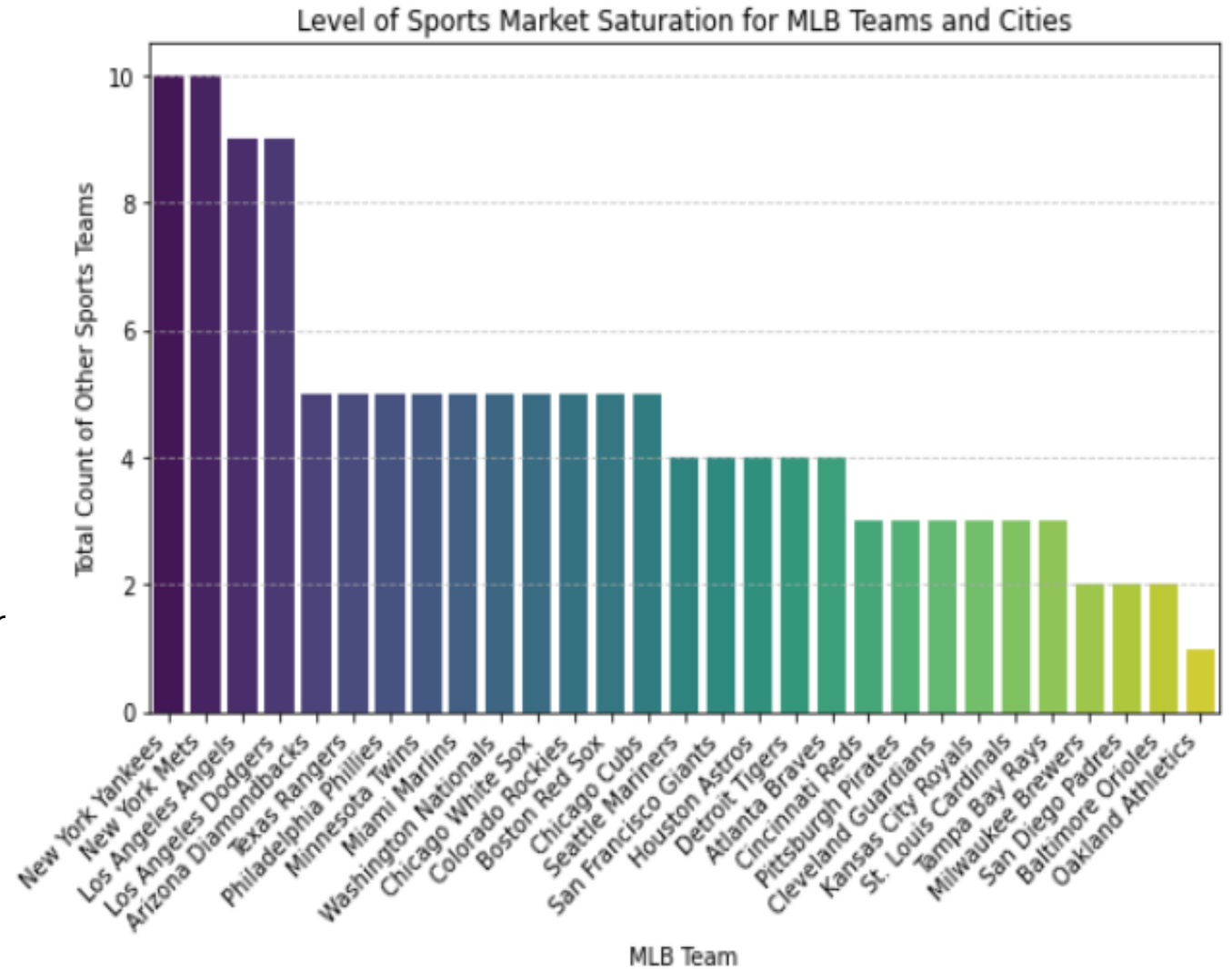
- A positive correlation between population and revenue is evident (e.g., New York and Los Angeles lead both population and revenue).
- Some smaller cities like San Francisco achieve high revenues despite lower populations, likely due to high median incomes or strong brand value compared to Houston and Arlington with larger Population.
- Income levels play a role, as higher-income cities (e.g., San Francisco, Boston) tend to generate higher revenues even with moderate populations.



Current Market Size of MLB Teams

Sports Market Saturation

- Cities like New York and Los Angeles are heavily saturated with other professional teams.
- This saturation may create competition for fan engagement but also indicates strong sports markets.
- Cities with fewer competing teams (e.g., Seattle or Oakland) may rely more heavily on MLB for sports entertainment.



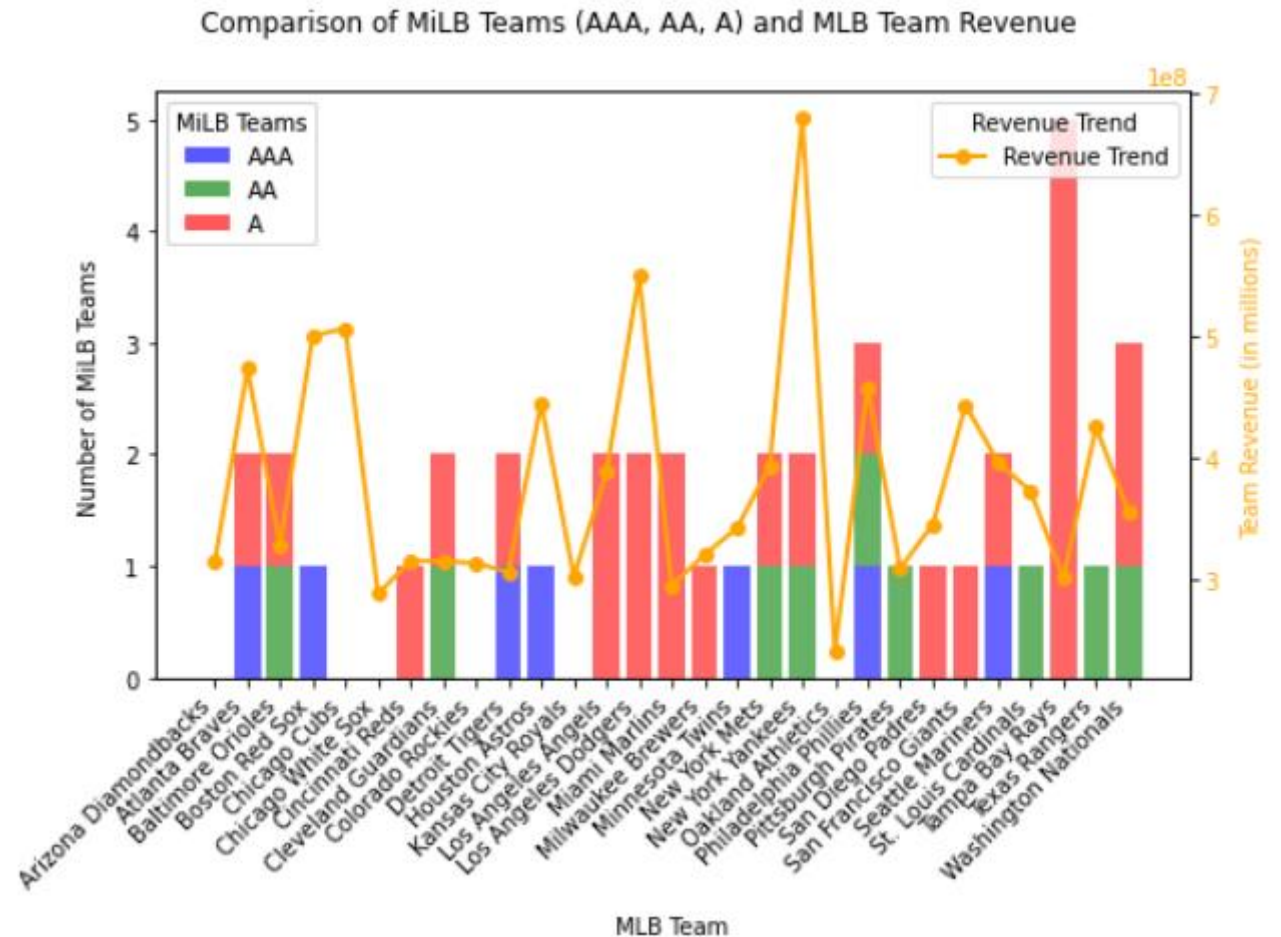
Current Market Size of MLB Teams

Relationship Between MiLB Teams and Revenue:

- Teams with a balanced presence across AAA, AA, and A levels in their Cities tend to show stronger ecosystems, which might correlate with consistent revenue (e.g., Yankees, Giants).
- High revenue teams don't always rely on a strong MiLB presence (e.g., Dodgers), suggesting market size and branding could be factors.

Outliers:

- Washington Nationals:** High AAA team presence but moderate revenue, highlighting a potential gap in leveraging minor league ecosystem presence in the city.
- Tampa Bay Rays:** Low MiLB presence in the city and low revenue,



FINDINGS

Revenue Drivers:

- Population and income are factors but are not absolute determinants.
- MiLB Presence Supports Revenue, But It's Not a primary driver:

Market Saturation:

- Saturated cities like New York and Los Angeles can still thrive due to their large populations and economic power.
- Markets with fewer competing teams may offer opportunities for deeper fan engagement.

Expansion Opportunities:

- New MLB markets should prioritize cities with a balance of population, income, and minimal competition from other sports like NFL, MLS, NHL, NBA and also Cities with MiLB (esp. AAA)



METHOD/APPROACH

- A regression model was used to determine the relationship between various factors (population, income, sports saturation and total Pro League) and MLB team revenue.
 - The model was trained using both standardized and unstandardized features to calculate coefficients and identify the importance of each predictor."
 - **Features included:**
 - Populations (city population and metropolitan population) and income metrics.
 - Counts of other professional sports teams (A new feature, total_Pro_league, counts of all professional sports teams in a city)
 - Minor league baseball presence (AAA, AA, A).
-

Features Importance and Impacts on MLB Revenue

Standardized Coefficient: compare the relative importance of features

- Larger absolute values indicate stronger effects on revenue.

Unstandardized Coefficient: the actual impact of each feature on revenue

P-Value:

- Indicates the statistical significance of the feature:

- **$p < 0.05$** : Statistically significant.

- **$p \geq 0.05$** : Not statistically significant.

Significance:

- Clearly labels each feature as 'Significant' or 'Not Significant' based on the p-value.

Features	Standardized Coefficient	Unstandardized Coefficient	P-Value	Significance
City Population Estimate	25840500	1.086862e+01	0.367946	NO
Median Household Income (City)	10282910	4.634968e+02	0.537526	NO
Metropolitan Population Estimate	17653830	3.816077e+00	0.689971	NO
NFL Count	923996.4	-7.229860e+06	0.410123	NO
NBA Count	14038120	1.270969e+07	0.581616	NO
NHL Count	-25795520	-4.048852e+07	0.030477	YES
MLS Count	33184880	4.566937e+07	0.780204	NO
AAA	30937830	7.229821e+07	0.056851	NO
AA	17571270	3.798088e+07	0.329891	NO
A	8691899	8.346709e+06	0.609195	NO
Total_Pro_league	3789407	1.066068e+07	0.004223	YES

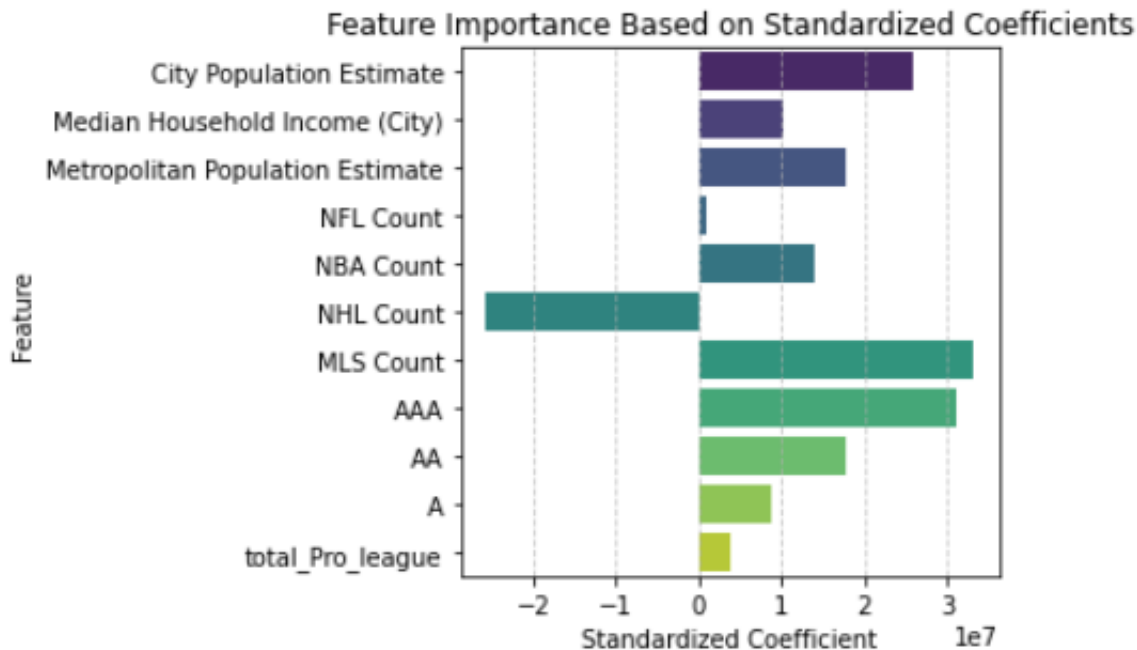
Regression Model Equation/Performance

MLB Revenue = $2.067105 \times 10^8 + (1.086862 \cdot \text{City Population Estimate}) + (463.4968 \cdot \text{Median Household Income (City)}) + (3.816077 \cdot \text{Metropolitan Population Estimate}) - (7.229860 \times 10^6 \cdot \text{NFL Count}) + (1.270969 \times 10^7 \cdot \text{NBA Count}) - (4.048852 \times 10^7 \cdot \text{NHL Count}) + (4.566937 \times 10^7 \cdot \text{MLS Count}) + (7.229821 \times 10^7 \cdot \text{AAA}) + (3.798088 \times 10^7 \cdot \text{AA}) + (8.346709 \times 10^6 \cdot \text{A}) + (1.066068 \times 10^7 \cdot \text{total_Pro_league})$

Model Performance

- R-squared: 0.59
 - The model explains approximately 59% of the variance in MLB team revenue.
 - Mean Squared Error (MSE): $\$3.6 \times 10^{15}$
 - The average squared difference between predicted and actual revenues.
-

Interpretations



Population as a Revenue Driver:

- **City Population Estimate** has the strongest positive influence on MLB revenue, underscoring the importance of targeting larger markets for MLB expansion or team success.

Economic Factors :

- While **Median Household Income (City)** shows a positive impact, it is not statistically significant, indicating that revenue depends more on population size than local income levels.

Professional Sports Competition (Market Saturation):

- **NHL Count** negatively impacts revenue the most among competing leagues (\$-40.49M per additional team), highlighting the challenges MLB faces in hockey-dominated markets.
- Conversely, **MLS Count** shows a positive (though not statistically significant) correlation, suggesting soccer may complement rather than compete with MLB in some cities.

Interpretations

Total Professional Teams as a Cultural Indicator:

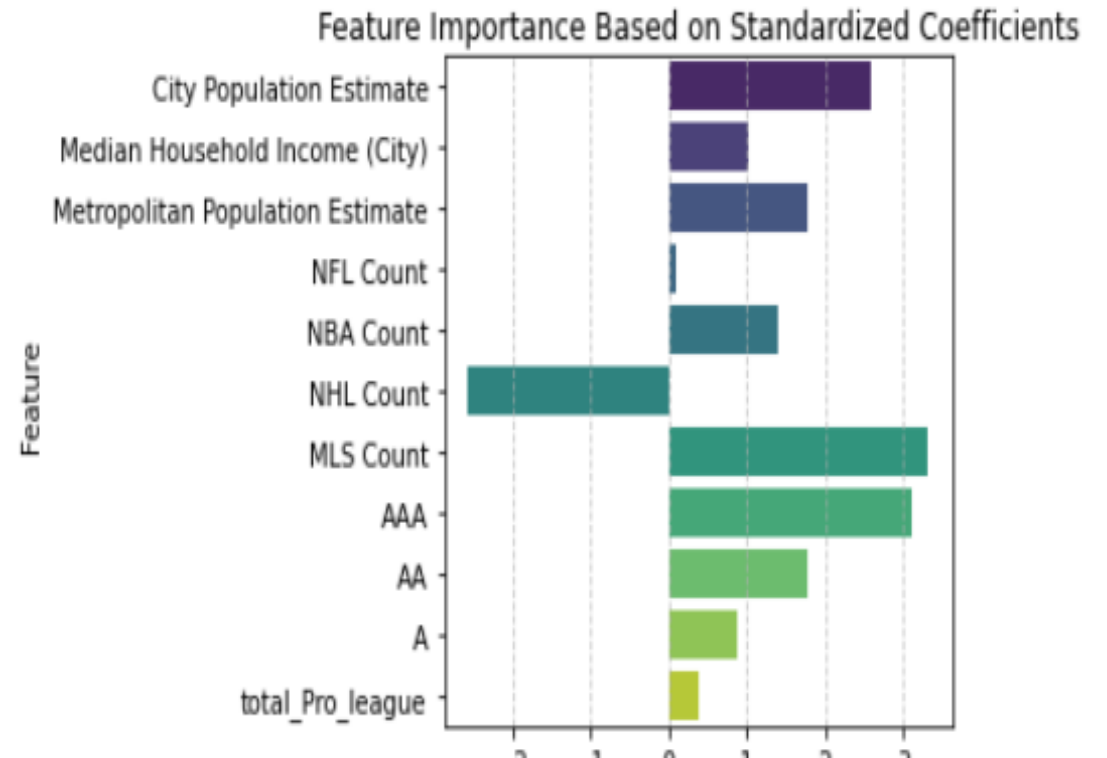
- Cities with a higher **total_Pro_league** (count of NFL, NBA, NHL, MLS, and MLB teams) demonstrate a positive revenue impact (\$10.66M per additional team), indicating that a vibrant sports culture benefits MLB.

Minor League Baseball (MiLB) Presence:

- AAA Teams** provide the strongest minor league contribution to revenue (\$72.30M per team), emphasizing the importance of top-tier minor league systems in developing fan engagement and talent pipeline.

Insights

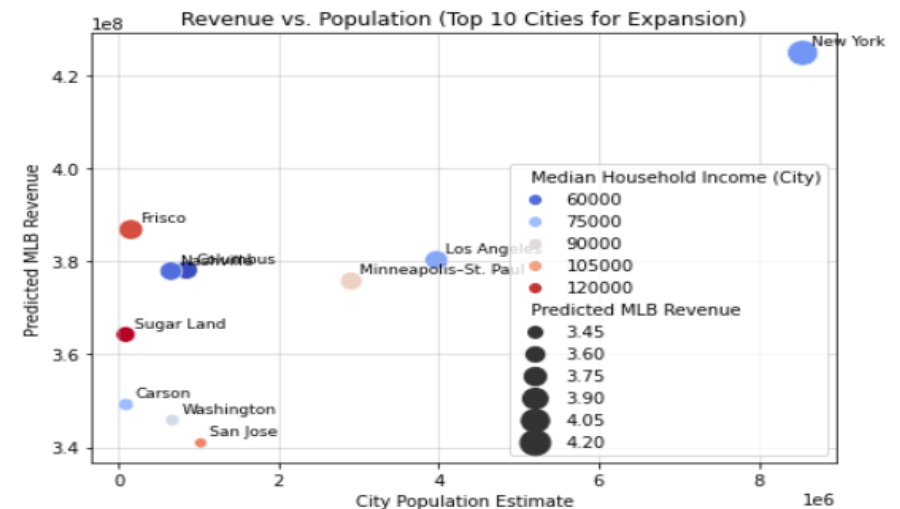
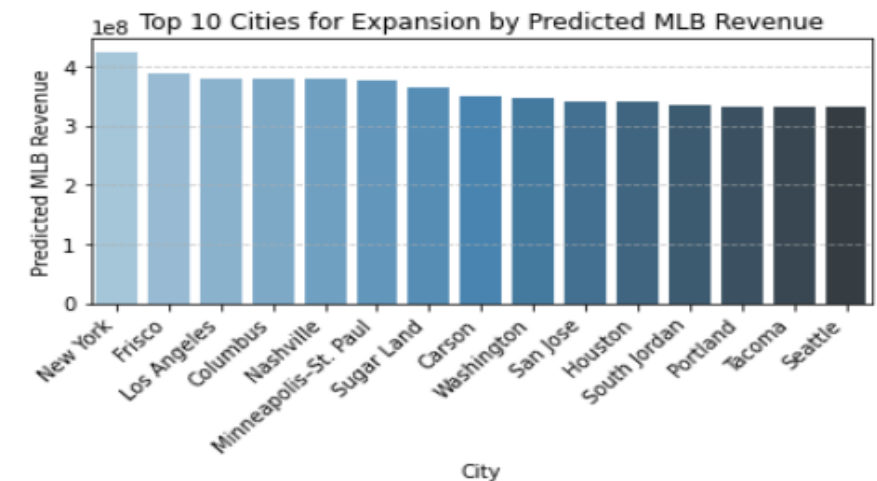
- Expansion into larger markets with established sports cultures and a strong minor league presence (especially AAA) is likely to maximize revenue potential.
- Cities with fewer NHL teams may offer less competition for MLB fan engagement and higher revenue opportunities.



TOP 10 Cities

Among the 385 cities selected :

- New York, Frisco, and Los Angeles dominate due to their combination of market size and/or economic strength.
- Cities like Carson, Washington, and San Jose rank lower, reflecting their more saturated markets or limited population bases.
- Smaller cities like Frisco and Sugar Land have lower populations but rank high in revenue due to higher median incomes.
- Los Angeles, Minneapolis–St. Paul, and Nashville cluster near the top, balancing population and income.



TOP CITIES WITH POSSIBLE REVENUE

New York Leads by a Wide Margin:

- Predicted MLB Revenue: \$424.96M.
- Population: 8.55M with moderate median income (\$67K).
- Despite already being a saturated market, New York's size and economic activity still make it the top candidate.

Frisco as a High Revenue Candidate:

- Predicted Revenue: \$386.89M, ranking 2nd.
 - Despite its smaller population (154K), Frisco benefits from a **high median household income (\$116.88K)**, indicating significant spending power in the region.
-

TOP CITIES WITH POSSIBLE REVENUE

Los Angeles Stands Out:

- With a **large population (3.97M)** and strong economic activity, Los Angeles ranks 3rd with **\$380.33M** in predicted revenue.
- However, this market may already be saturated with other professional sports teams.

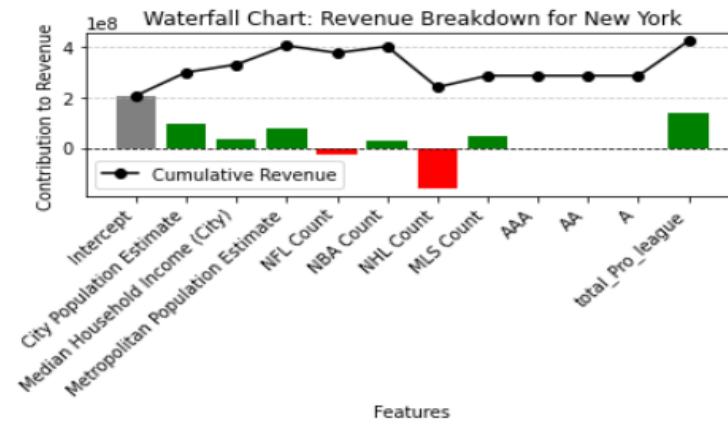
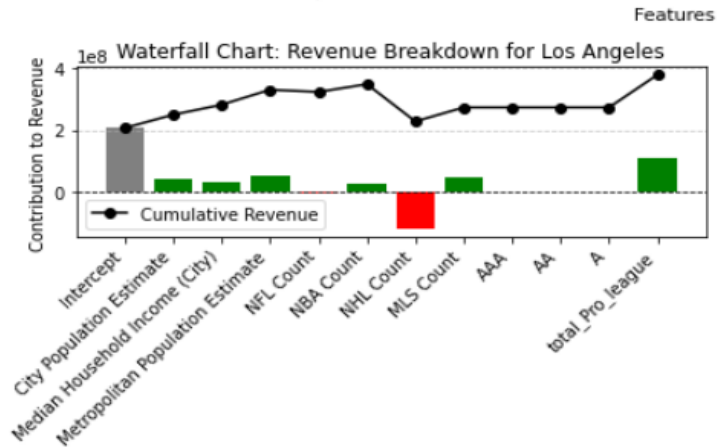
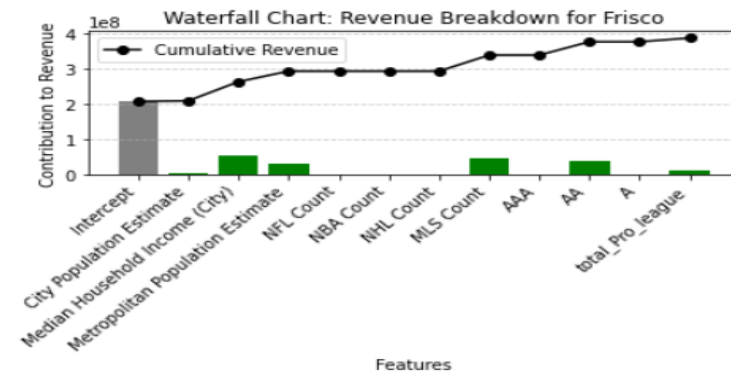
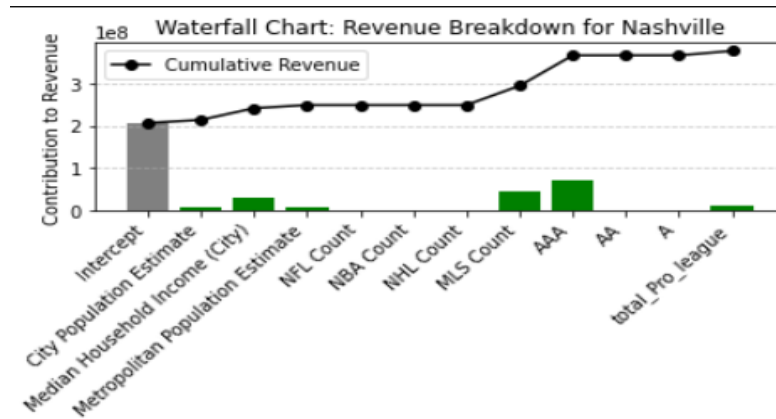
Cities with Balanced Growth Potential:

- **Columbus (\$378.18M)**, **Nashville (\$377.94M)**, and **Minneapolis–St. Paul (\$375.81M)** rank 4th, 5th, and 6th, respectively, combining **moderate population and income levels** with potential for additional sports market penetration.

Smaller Markets with High Income:

- **Sugar Land (\$364.28M)**: Although the population is small (88K), its high median income (\$123.26K) elevates its potential revenue.
-

TOP CITIES AND CONTRIBUTOR FACTORS



Recommendations

Prioritize Large Cities with Untapped Potential

- **Nashville, Columbus, and Minneapolis–St. Paul** provide strong growth opportunities with moderate population sizes and spending power.
- These cities also face less saturation compared to **New York and Los Angeles**.

Consider High-Income, Smaller Markets

- Cities like **Frisco and Sugar Land** highlight the potential of affluent **but smaller markets**. Their high **spending power compensates** for their lower **population bases**.

Account for Market Saturation:

- While cities like **New York and Los Angeles** generate **high predicted revenues**, their existing **sports saturation** may limit opportunities for additional MLB teams.
-

ADDITIONAL METRICS

- These metrics would also be helpful for further analysis
- Fan Engagement Metrics
- Sponsorship and Partnership Deals

REFERENCES

- <https://www.forbes.com/lists/mlb-valuations/>
 - <https://www.census.gov/>
 - https://infogalactic.com/info/U.S._cities_with_teams_from_four_major_league_sports
-