Analysis Memo #1

Checkpoint 6

Mapping Global Growth — A Data-Driven Strategy for NFL Market Entry

Question and Outcome

Question: Which international markets present the highest potential for sustainable NFL growth?

Objective: Understand which countries could support the NFL's international expansion in terms of fan engagement, event attendance, and long-term market development. Insights from this analysis help prioritize international marketing and expansion strategies.

Outcome Column: For this analysis, the primary outcome metric is a composite "market potential score", calculated using indicators such as stadium capacity, connectivity index, population size, average income, and historical sports engagement. This score serves as a proxy for overall market attractiveness and growth potential.

Data Used

The project is based on the cleaned country-level dataset from CP4 (merged_data), where each row represents a single country. Variables include Google Trends signals (to measure interest in the NFL), economic indicators (GDP per capita), demographics (urban population), stadium infrastructure, and aviation connectivity metrics.

The following countries were dropped due to either 1. Already hosting NFL games or 2. Not having complete information across all metrics:

1. United States, United Kingdom, Ireland, Spain, Germany, Brazil

2. 'American Samoa', 'Andorra', 'Aruba', 'Belize', 'Bermuda', 'Bhutan', 'British Virgin Islands', 'Cayman Islands', 'Comoros', 'Dominica', 'Eritrea', 'Eswatini', 'Faroe Islands', 'French Polynesia', 'Gibraltar', 'Greenland', 'Grenada', 'Guam', 'Guyana', 'Haiti', 'Iceland', 'Isle of Man', 'Kiribati', 'Kosovo', 'Latvia', 'Lebanon', 'Liechtenstein', 'Luxembourg', 'Marshall Islands', 'Mongolia', 'Namibia', 'Nauru', 'New Caledonia', 'Northern Mariana Islands', 'Palau', 'Samoa', 'San Marino', 'Sierra Leone', 'Solomon Islands', 'South Sudan', 'St. Lucia', 'Suriname', 'Switzerland', 'Timor-Leste', 'Tonga', 'Tuvalu', 'Vanuatu'

Method Choice

1. **Market Attractiveness Index:**
   A weighted composite index combining the key variables: NFL engagement, economic readiness, infrastructure readiness, and connectivity index. Each variable was normalized to ensure comparability before weighting. This index provides a single, interpretable metric that ranks countries by overall market attractiveness. The rankings may vary according to the weights attributed to each aspect of the index.

   Baseline MAI (equally distributed weights):

   $$MAI = 0.25 * \text{fan engagement} + 0.25 * \text{economic readiness} + 0.25 * \text{infrastructure readiness} + 0.25 * \text{connectivity index}$$

2. **K-Means Clustering:**
   To go beyond a simple ranking, k-means clustering was also applied to the same set of normalized variables. This method groups countries with similar characteristics, allowing us to identify distinct market segments that share common traits.

Rationale:

- The composite index provides a straightforward ranking of countries, which is easy to communicate and serves as a baseline. Can be further tweaked, based on different rationales for each of the weights (e.g., prioritizing economic/infrastructure readiness over fan engagement, or vice-versa).

- Clustering adds depth by revealing patterns across multiple variables simultaneously, which a table or chart alone cannot show. It helps identify which markets are similar and may respond similarly to NFL expansion strategies.

Analysis Spec

The following table summarizes the main elements of the analysis:

Element	Specification
Outcome	Composite Market Potential Index (weighted, normalized score per country)
Predictors / Groups	Stadium capacity, connectivity index, population, GDP per capita, NFL fan engagement (from Google trends)
Sample	128 International countries included in the dataset (excluding extreme outliers where noted and missing values)
Row Definition	One row per country
Rule / Formula	Standardize all input variables, apply predefined weights,sum weighted values to produce composite score. K-means applied to normalized inputs to define market clusters.

Inputs Included and Rationale

- **Infrastructure Readiness**: Proxy for a country's ability to host large live events.
- **Connectivity Index**: Measures how accessible the country is.
- **Economic Readiness**: Proxy for purchasing power and ability to support premium sports entertainment.
- **Fan Engagement**: Measures cultural openness/interest for NFL events.

Expected Direction of Effects: All variables should move in the same direction as the index (higher values of the variable, higher MAI score).

For clustering, the expectation is that countries with high infrastructure + high economic measures to group together, and countries with large populations but weaker infrastructure to form distinct clusters.

The number of clusters was selected according to the 'Elbow Plot':

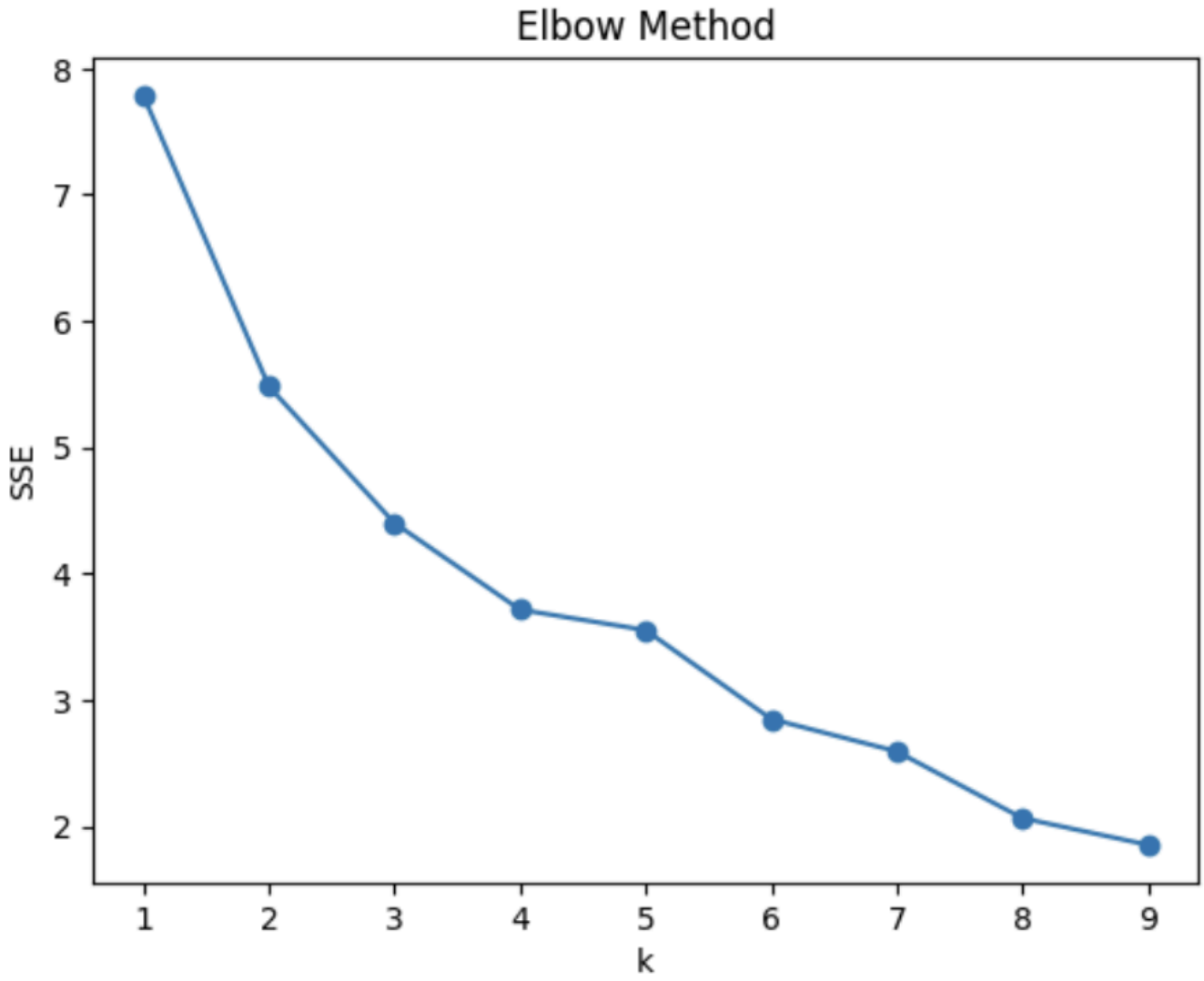

Fig 1. Clustering Elbow Plot - Selected 4 as the optimal k.

Results

1. **Top Markets by Market Attractiveness Index (MAI)**

Rank	Country	MAI
1	Canada	0.596

2	China	0.518
3	France	0.435
4	Mexico	0.421
5	Australia	0.400
6	Netherlands	0.372
7	Qatar	0.363
8	United Arab Emirates	0.340
9	Austria	0.340
10	Singapore	0.334

The spread in MAI scores is quite large: the top market (Canada, 0.596) scores roughly 78% higher than the lowest country in the top 10 (Singapore, 0.334). This range suggests substantial differentiation in market readiness rather than small, noise-level differences.

2. **Cluster Summary**

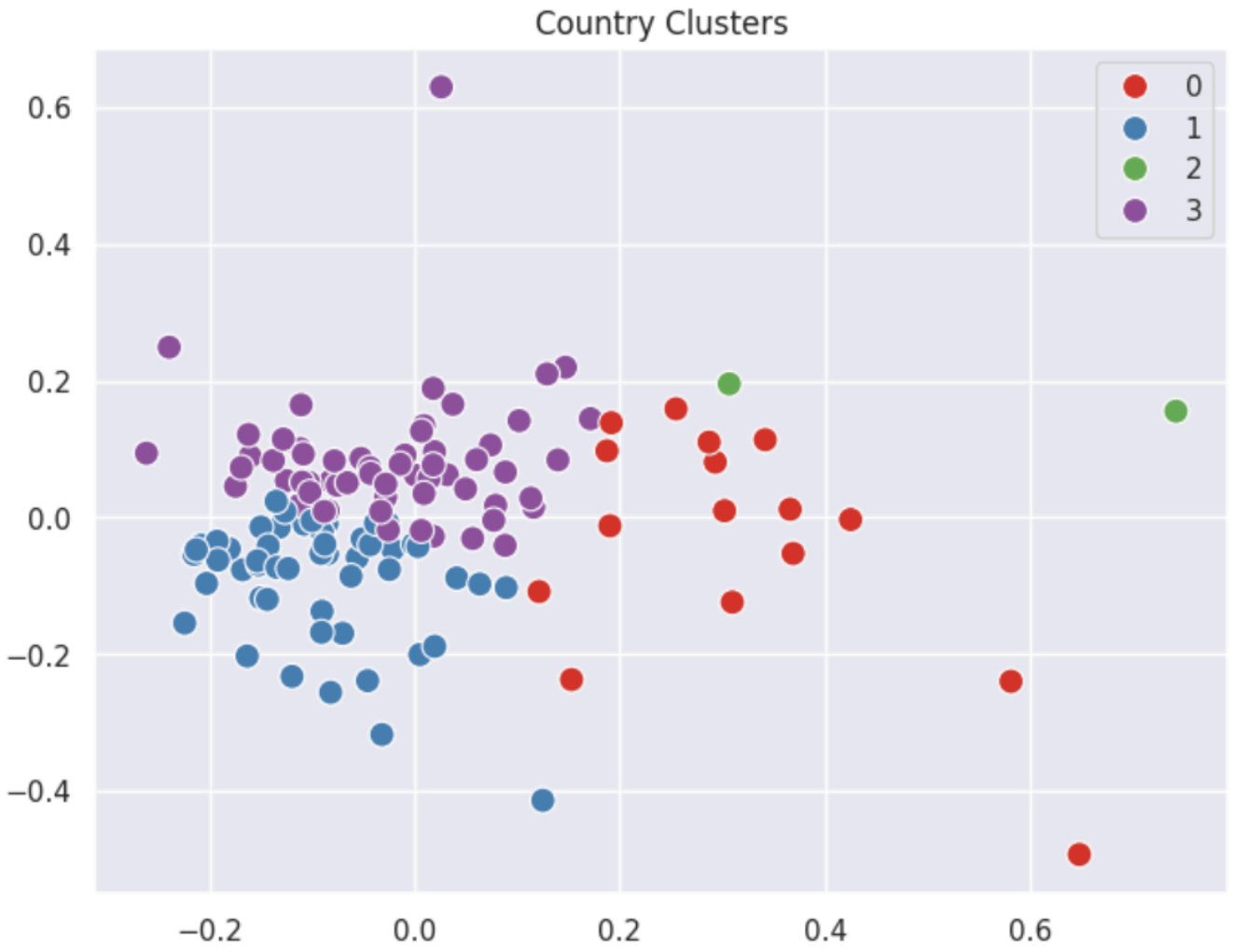

Fig 2. Clustering outcome (after PCA) - Country clusters based on fan engagement, economic readiness, infrastructure readiness, and connectivity. Each point represents a country projected into two principal components, with colors indicating k-means cluster membership.

Cluster	Profile	Examples of Countries
Cluster 0	**High-readiness, infrastructure-strong markets**	Australia, China, France, Japan, Netherlands, Singapore, UAE, Qatar
Cluster 1	**Low readiness / emerging infrastructure markets**	Nigeria, Kenya, Bangladesh, Vietnam, Pakistan, Ethiopia, Nepal

Cluster 2	**Top tier North American–style markets**	Canada, Mexico
Cluster 3	**Mid-readiness, mixed infrastructure markets**	Argentina, Poland, Portugal, Saudi Arabia, Malaysia, Chile, Colombia

The clusters reveal the following market tiers:

- Cluster 0 represents immediate expansion candidates with strong infrastructure and economic fundamentals.
- Cluster 2 (Canada and Mexico) behaves as a separate high-potential group, reflecting geographic proximity and existing cultural exposure to American football.
- Cluster 3 contains promising development markets that could support medium-term growth strategies.
- Cluster 1 represents countries less suited for the NFL's international expansion.

Silhouette score of 0.285 is relatively low, suggesting that there can be improvements to the method.

Checks

Input Scale Validation
All input variables (fan engagement, economic readiness, infrastructure readiness, and connectivity) were normalized before building the Market Attractiveness Index (MAI). This ensured that no single variable dominated the results purely due to scale differences.

Data Check
All variables reflect **structural market characteristics** (infrastructure, population, economic indicators, and engagement proxies) rather than future outcomes. No forward-looking or post-expansion data was used, reducing the risk of look-ahead bias.

Outlier and Sample Sensitivity
The stability of the rankings was tested by repeatedly recalculating the MAI using 300 random weight combinations and tracking how often each country appeared in the top 10. This simulates uncertainty in how strongly each input should be weighted and tests whether results are overly dependent on a specific weighting scheme.

Stability Test Results

Country	Top-10 Appearance Rate
Canada	97.3%
China	93.3%
Australia	90.7%
France	87.0%
Netherlands	77.0%
Qatar	72.0%
Mexico	62.7%
Austria	49.7%
United Arab Emirates	40.3%
Singapore	38.0%

Interpretation
The top markets (Canada, China, Australia, and France) appear in the top 10 in over 85% of simulations, indicating that these results are highly robust to reasonable changes in model assumptions. Lower-ranked countries show more variability, suggesting they are more sensitive to weighting choices and should be viewed as borderline or secondary expansion candidates.

Interpretation

MAI
Countries at the top of the list are more likely to support regular-season games, strong broadcast demand, and sustained fan growth, while lower-ranked countries in the top 10 may be better suited for pilot events or long-term brand-building rather than immediate large-scale expansion.

Clustering

- Clusters 0 and 2 seem to be realistic candidates for regular-season international games.
- Cluster 3 could support preseason games, fan festivals, and broadcast-first strategies.
- Cluster 1 is better suited for revisiting further down the road.

Limitations

Simplified Market Proxies
This analysis relies on proxy variables (infrastructure, economic readiness, connectivity, and sports engagement) rather than direct NFL-specific demand signals such as TV ratings for American football, merchandise sales, or fantasy sports participation. As a result, the model may overestimate market attractiveness in countries with strong general infrastructure but limited NFL-specific interest.

Subjective Weighting and Feature Gaps
The composite index depends on assumed variable weights and a limited feature set. Important factors such as time zone alignment with U.S. broadcasts, competitive competition from dominant local sports leagues, and historical attendance for sporting events were not included, which may bias results.

Next Steps

- Test alternative clustering methods (e.g., hierarchical clustering) to validate segmentation stability and improve Silhouette Score.
- Run a more robust sensitivity analysis with structured weight scenarios instead of purely random weights.