```
taruns@Taruns-MacBook-Air-2 Plastic-Waste-Flow-Recycling-Prediction-main % python3 app.py
================================================================================
PLASTIC WASTE FLOW & RECYCLING PREDICTION - COMPLETE ANALYSIS
================================================================================
Team: Tarun S, Adityaa Kumar H, Akshay P Shetti
Course: Advanced Data Analytics (CSE-AIML) UE23AM343AB1
================================================================================
[STEP 1/6] LOADING DATA
--------------------------------------------------------------------------------
======================================================================
LOADING ALL DATASETS
======================================================================
✓ Loaded production data: 69 records
  Year range: 1950 - 2019
  Production range: 2,000,000 - 459,746,020 tonnes
✓ Loaded mismanaged waste data: 165 countries/regions
  Year: [2019]
  Waste per capita range: 0.00 - 69.52 kg/year
✓ Loaded trade data (sample): 15644 records (50% of total)
  Columns: 48
  Year range: 2012 - 2023
  Unique reporters: 179
  Unique partners: 226
======================================================================
ALL DATASETS LOADED SUCCESSFULLY
======================================================================
Dataset Summary:
PRODUCTION:
  total_records: 69
  year_range: (1950, 2019)
  production_range_tonnes: (2000000, 459746020)
  avg_annual_growth_rate: 8.20
MISMANAGED_WASTE:
  total_countries: 165
  year: 2,019
  waste_range_kg: (0.002719679, 69.51586369)
  global_avg_kg: 8.01
  top_5_countries: [{'Entity': 'Comoros', 'Waste_Per_Capita_kg': 69.51586369}, {'Entity': 'Trinidad and Tobago',
'Waste_Per_Capita_kg': 52.42939068}, {'Entity': 'Suriname', 'Waste_Per_Capita_kg': 39.47160069}, {'Entity':
'Philippines', 'Waste_Per_Capita_kg': 37.23096275}, {'Entity': 'Zimbabwe', 'Waste_Per_Capita_kg': 35.83919426}]
TRADE:
  total_records: 15,644
  year_range: (2012, 2023)
  unique_reporters: 179
  unique_partners: 226
  total_quantity_kg: 116535728437.26
  total_value_usd: 54403989631.40
[STEP 2/6] PREPROCESSING DATA
--------------------------------------------------------------------------------
======================================================================
PREPROCESSING ALL DATASETS
======================================================================
Preprocessing production data...
  ✓ Cleaned 69 records
  ✓ Added 4 derived features
Preprocessing mismanaged waste data...
  ✓ Cleaned 159 countries
  ✓ Removed 6 regional aggregates
  ✓ Added categorization and ranking
Preprocessing trade flow data...
  ✓ Cleaned 14181 trade records
  ✓ Selected 14 key columns
```

```
  ✓ Removed outliers and missing values
Creating network data structure...
  ✓ Created 4693 edges
  ✓ Unique nodes (countries): 226
Aggregating trade data by year...
  ✓ Aggregated into 12 years
======================================================================
PREPROCESSING COMPLETE
======================================================================
Saving preprocessed data to outputs/...
  ✓ Saved production: 69 records
  ✓ Saved waste_countries: 159 records
  ✓ Saved waste_world: 1 records
  ✓ Saved trade: 14181 records
  ✓ Saved trade_network: 4693 records
  ✓ Saved trade_yearly: 12 records
All preprocessed data saved successfully!
[STEP 3/6] FLOW ANALYSIS
--------------------------------------------------------------------------------
Creating trade network graph...
Using minimum weight filter: 1000 tonnes
Creating network graph for Export flows...
  Applied minimum weight filter: 1000 tonnes
  ✓ Created graph with 145 nodes and 1405 edges
Calculating network metrics...
Calculating centrality metrics...
  Computing degree centrality...
  Computing betweenness centrality...
  Computing closeness centrality...
  Computing PageRank...
  Computing eigenvector centrality...
  ✓ Centrality metrics calculated
--- Top 10 Countries by PageRank ---
 1. China                    0.2273
 2. USA                      0.0715
 3. China, Hong Kong SAR     0.0596
 4. Malaysia                 0.0587
 5. India                    0.0303
 6. Viet Nam                 0.0232
 7. Thailand                 0.0205
 8. South Africa             0.0168
 9. Singapore                0.0159
10. Canada                   0.0156
--- Top 10 Countries by Betweenness Centrality ---
 1. China                    0.1646
 2. USA                      0.1091
 3. Viet Nam                 0.0945
 4. Belgium                  0.0857
 5. Greece                   0.0613
 6. South Africa             0.0602
 7. India                    0.0595
 8. Spain                    0.0581
 9. Poland                   0.0510
10. China, Hong Kong SAR     0.0497
Calculating network statistics...
Calculating network statistics...
  ✓ Network statistics calculated
    num_nodes: 145
    num_edges: 1405
    density: 0.06728927203065134
    is_connected: True
    num_connected_components: 1
```

```
   avg_degree: 19.379310344827587
   max_degree: 130
   min_degree: 1
   avg_shortest_path_length: None
   avg_clustering_coefficient: 0.5808296562629731
   total_flow_tonnes: 62780467.7829
Identifying critical nodes...
Identifying 10 critical nodes...
  ✓ Critical nodes identified:
   1. China: 0.1646
   2. USA: 0.1091
   3. Viet Nam: 0.0945
   4. Belgium: 0.0857
   5. Greece: 0.0613
   6. South Africa: 0.0602
   7. India: 0.0595
   8. Spain: 0.0581
   9. Poland: 0.0510
   10. China, Hong Kong SAR: 0.0497
[STEP 4/6] TIME SERIES FORECASTING
--------------------------------------------------------------------------------
--- Stationarity Test ---
ADF Statistic: 6.6395
p-value: 1.0000
Is Stationary: False
--- ARIMA Forecasting ---
============================================================
ARIMA Forecasting (order=(2, 1, 2))
============================================================
Fitting ARIMA(2, 1, 2) model...
✓ Model fitted successfully
Model Summary:
                               SARIMAX Results
==============================================================================
Dep. Variable:      Production_Tonnes   No. Observations:                   69
Model:                 ARIMA(2, 1, 2)   Log Likelihood               -1144.706
Date:                Tue, 18 Nov 2025   AIC                           2299.411
Time:                        09:39:46   BIC                           2310.509
Sample:                             0   HQIC                          2303.808
                                  - 69
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.3566      0.251      1.421      0.155      -0.135       0.848
ar.L2          0.6434      0.285      2.261      0.024       0.086       1.201
ma.L1         -0.0010      0.206     -0.005      0.996      -0.404       0.402
ma.L2         -0.6622      0.218     -3.035      0.002      -1.090      -0.235
sigma2       2.62e+13   2.22e-14   1.18e+27      0.000    2.62e+13    2.62e+13
===================================================================================
Ljung-Box (L1) (Q):                   2.45   Jarque-Bera (JB):               365.09
Prob(Q):                              0.12   Prob(JB):                         0.00
Heteroskedasticity (H):              31.23   Skew:                            -1.78
Prob(H) (two-sided):                  0.00   Kurtosis:                        13.78
===================================================================================
Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
[2] Covariance matrix is singular or near-singular, with condition number 4.26e+42. Standard errors may be unstable.
✓ Forecast for next 10 periods:
          Forecast  Model
Year
2020  4.758895e+08  ARIMA
```

2021 4.926510e+08 ARIMA
2022 5.090149e+08 ARIMA
2023 5.256346e+08 ARIMA
2024 5.420897e+08 ARIMA
2025 5.586507e+08 ARIMA
2026 5.751436e+08 ARIMA
2027 5.916803e+08 ARIMA
2028 6.081888e+08 ARIMA
2029 6.247155e+08 ARIMA
✓ ARIMA forecast completed
AIC: 2299.41
BIC: 2310.51
--- Prophet Forecasting ---
==
Prophet Forecasting
==
Fitting Prophet model...
09:39:46 - cmdstanpy - INFO - Chain [1] start processing
09:39:46 - cmdstanpy - INFO - Chain [1] done processing
✓ Model fitted successfully
✓ Forecast for next 10 periods:

	Year	Forecast	yhat_lower	yhat_upper
69	2019	4.294276e+08	4.224299e+08	4.370464e+08
70	2020	4.612137e+08	4.538251e+08	4.685603e+08
71	2021	4.695333e+08	4.616082e+08	4.776012e+08
72	2022	4.762524e+08	4.678183e+08	4.843198e+08
73	2023	4.813814e+08	4.729959e+08	4.896779e+08
74	2024	5.131676e+08	5.041334e+08	5.216708e+08
75	2025	5.214872e+08	5.122683e+08	5.306791e+08
76	2026	5.282063e+08	5.184102e+08	5.380752e+08
77	2027	5.333352e+08	5.230879e+08	5.442565e+08
78	2028	5.651214e+08	5.540910e+08	5.767601e+08

✓ Prophet forecast completed
--- LSTM Forecasting ---
==
LSTM Forecasting
==
Preparing data with lookback=5...
Training samples: 51, Test samples: 13
Building LSTM model...
Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 5, 50)	10,400
dropout (Dropout)	(None, 5, 50)	0
lstm_1 (LSTM)	(None, 50)	20,200
dropout_1 (Dropout)	(None, 50)	0
dense (Dense)	(None, 1)	51

Total params: 30,651 (119.73 KB)

Trainable params: 30,651 (119.73 KB)
Non-trainable params: 0 (0.00 B)
None
Training for 50 epochs...
✓ Training complete
Final training loss: 0.001476
Final validation loss: 0.003091
Generating 10 period forecast...
✓ Forecast for next 10 periods:

Year	Forecast	Model
2020	5.412212e+08	LSTM
2021	5.896599e+08	LSTM
2022	6.557622e+08	LSTM
2023	7.495954e+08	LSTM
2024	8.898856e+08	LSTM
2025	1.118733e+09	LSTM
2026	1.450015e+09	LSTM
2027	2.076151e+09	LSTM
2028	3.435258e+09	LSTM
2029	6.803766e+09	LSTM

✓ LSTM forecast completed
--- Model Comparison ---
==
MODEL COMPARISON
==

Year	arima	prophet	lstm	Mean	Std
2020	4.758895e+08	4.294276e+08	5.412212e+08	4.821794e+08	4.585576e+07
2021	4.926510e+08	4.612137e+08	5.896599e+08	5.145082e+08	5.466814e+07
2022	5.090149e+08	4.695333e+08	6.557622e+08	5.447701e+08	8.012129e+07
2023	5.256346e+08	4.762524e+08	7.495954e+08	5.838275e+08	1.189367e+08
2024	5.420897e+08	4.813814e+08	8.898856e+08	6.377856e+08	1.799763e+08
2025	5.586507e+08	5.131676e+08	1.118733e+09	7.301839e+08	2.753727e+08
2026	5.751436e+08	5.214872e+08	1.450015e+09	8.488820e+08	4.256294e+08
2027	5.916803e+08	5.282063e+08	2.076151e+09	1.065346e+09	7.152165e+08
2028	6.081888e+08	5.333352e+08	3.435258e+09	1.525594e+09	1.350682e+09
2029	6.247155e+08	5.651214e+08	6.803766e+09	2.664534e+09	2.926980e+09

Year	arima	prophet	lstm	Mean	Std
2020	4.758895e+08	4.294276e+08	5.412212e+08	4.821794e+08	4.585576e+07
2021	4.926510e+08	4.612137e+08	5.896599e+08	5.145082e+08	5.466814e+07
2022	5.090149e+08	4.695333e+08	6.557622e+08	5.447701e+08	8.012129e+07
2023	5.256346e+08	4.762524e+08	7.495954e+08	5.838275e+08	1.189367e+08
2024	5.420897e+08	4.813814e+08	8.898856e+08	6.377856e+08	1.799763e+08
2025	5.586507e+08	5.131676e+08	1.118733e+09	7.301839e+08	2.753727e+08
2026	5.751436e+08	5.214872e+08	1.450015e+09	8.488820e+08	4.256294e+08
2027	5.916803e+08	5.282063e+08	2.076151e+09	1.065346e+09	7.152165e+08
2028	6.081888e+08	5.333352e+08	3.435258e+09	1.525594e+09	1.350682e+09
2029	6.247155e+08	5.651214e+08	6.803766e+09	2.664534e+09	2.926980e+09

[STEP 5/6] CLUSTERING & ANOMALY DETECTION

--- K-Means Clustering ---
==
K-Means Clustering (k=5)
==
Fitting K-Means with 5 clusters...
✓ Clustering complete
Cluster sizes:
Cluster 0: 76 countries (47.8%)
Cluster 1: 13 countries (8.2%)
Cluster 2: 42 countries (26.4%)

```
  Cluster 3: 2 countries (1.3%)
  Cluster 4: 26 countries (16.4%)
Cluster centers:
   Waste_Per_Capita_kg  Cluster
0             1.008104        0
1            29.027349        1
2             8.814529        2
3            60.972627        3
4            16.528470        4
Cluster Centers:
   Waste_Per_Capita_kg  Cluster
0             1.008104        0
1            29.027349        1
2             8.814529        2
3            60.972627        3
4            16.528470        4
--- Hierarchical Clustering ---
============================================================
Hierarchical Clustering (n=5, method=ward)
============================================================
Fitting hierarchical clustering...
✓ Clustering complete
Cluster sizes:
  Cluster 0: 13 countries (8.2%)
  Cluster 1: 42 countries (26.4%)
  Cluster 2: 76 countries (47.8%)
  Cluster 3: 2 countries (1.3%)
  Cluster 4: 26 countries (16.4%)
--- Isolation Forest Anomaly Detection ---
============================================================
Isolation Forest Anomaly Detection
============================================================
Fitting Isolation Forest (contamination=0.1)...
✓ Detection complete
  Anomalies detected: 16 (10.06%)
Top 10 anomalies:
                Entity  Waste_Per_Capita_kg  Anomaly_Score_IF
27             Comoros            69.515864         -0.846122
151  Trinidad and Tobago            52.429391         -0.774829
143           Suriname            39.471601         -0.688171
116        Philippines            37.230963         -0.643867
61              Guyana            35.204342         -0.635743
147           Tanzania            29.590553         -0.631485
164           Zimbabwe            35.839194         -0.616134
84               Libya            27.819832         -0.597718
158           Uruguay            26.753322         -0.572941
21           Cameroon            22.368140         -0.543611
--- Statistical Outlier Detection ---
============================================================
Statistical Outlier Detection (Z-score > 3)
============================================================
✓ Detection complete
  Outliers detected: 2 (1.26%)
Outliers:
                Entity  Waste_Per_Capita_kg  Max_Z_score
27             Comoros            69.515864     5.730148
151  Trinidad and Tobago            52.429391     4.121476
--- Trade Flow Anomalies ---
============================================================
Trade Flow Anomaly Detection
============================================================
✓ Detection complete
```

Threshold (P95.0): 13,711.82 tonnes
Anomalous flows: 709
Top 10 largest flows:

	Reporter	Partner	Year	Quantity_Tonnes	Value_USD
905	China, Hong Kong SAR	China	2012	3.160473e+06	1.008721e+09
8622	China, Hong Kong SAR	China	2015	2.797668e+06	8.850445e+08
13950	China, Hong Kong SAR	China	2017	1.710971e+06	5.407689e+08
5047	USA	China	2013	1.021279e+06	4.150288e+08
6271	Japan	China	2014	9.525344e+05	4.458810e+08
8862	Japan	China	2015	8.833773e+05	3.554995e+08
14204	Japan	China	2017	7.492677e+05	2.896438e+08
5888	Germany	China	2014	6.442833e+05	3.143892e+08
3318	Germany	China	2013	6.409186e+05	3.066779e+08
7648	USA	China, Hong Kong SAR	2014	5.659962e+05	2.190358e+08

Unusual pricing detected: 275 flows
[STEP 6/6] GENERATING VISUALIZATIONS

Creating visualizations...
✓ Saved: production_trend.html
✓ Saved: production_growth.html
✓ Saved: waste_choropleth.html
✓ Saved: top_countries_waste.html
✓ Saved: trade_network.html
✓ Saved: trade_flows_time.html
✓ Saved: clustering_results.html
✓ Saved: anomalies.html
✓ Saved: forecast_comparison.html
✓ Saved: summary_dashboard.html
✓ All visualizations saved to outputs/figures/
==========
ANALYSIS COMPLETE
==========
📊 KEY FINDINGS:
---
1. PRODUCTION TRENDS:
  • Global production grew from 2M (1950) to 460M tonnes (2019)
  • Average annual growth rate: ~8.5%
  • Exponential growth accelerated post-1990
2. WASTE MANAGEMENT HOTSPOTS:
  • Top 5 countries by mismanaged waste per capita:
    - Comoros: 69.52 kg/year
    - Trinidad and Tobago: 52.43 kg/year
    - Suriname: 39.47 kg/year
    - Philippines: 37.23 kg/year
    - Zimbabwe: 35.84 kg/year
3. TRADE NETWORK:
  • 145 countries involved in trade
  • 1405 trade connections
  • Network density: 0.0673
  • Critical hub countries identified
4. FORECASTS:
  • Multiple models predict continued growth
  • Production expected to reach ~600M tonnes by 2030
  • Urgent need for recycling infrastructure expansion
5. ANOMALIES:
  • 16 countries identified with anomalous waste patterns
  • Small island nations particularly vulnerable
📁 OUTPUT FILES:

  • Processed data: outputs/processed_*.csv
  • Visualizations: outputs/figures/*.html
  • Dashboard: Run 'python dashboard/app.py' to start

✅ All analysis complete!
  See outputs/ directory for results
  See README.md for detailed documentation
  See final_report.md for insights and recommendations
🚀 Main Dashboard:
  Click here to OPEN DASHBOARD
  (If the link above isn't clickable, copy this path into your browser):

file:///Users/taruns/Downloads/Plastic-Waste-Flow-Recycling-Prediction-main/outputs/figures/summary_dashboard.html
==========