Time Series Regression Model Development Report

# 1. Objective

The objective of this analysis is to project revenues for various businesses using a time series regression model that incorporates GDP as an external variable. For each business, a time series regression model is developed to understand the relationship between revenue and GDP and forecast future revenues.

# 2. Data Overview

The dataset used in this analysis includes the following columns:

|  |  |
| --- | --- |
| Column Name | Description |
| Date | Date of the record |
| Business | Business identifier |
| Revenue | Revenue for the business at the given date |
| GDP | Gross Domestic Product (GDP) at the given date |

Summary metrics for each business are as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Business | num\_records | start\_date | end\_date | avg\_revenue |
| biz A | 120 | 2013-01-31 00:00:00 | 2022-12-31 00:00:00 | 135.67084666666668 |
| biz B | 120 | 2013-01-31 00:00:00 | 2022-12-31 00:00:00 | 225.05726166666668 |
| biz C | 120 | 2013-01-31 00:00:00 | 2022-12-31 00:00:00 | 182.6973275 |
| biz D | 120 | 2013-01-31 00:00:00 | 2022-12-31 00:00:00 | 159.86244166666665 |
| biz E | 120 | 2013-01-31 00:00:00 | 2022-12-31 00:00:00 | 84.66819166666666 |

The following visualizations provide insights into revenue trends and their relationship with GDP for each business.

# 3. Exploratory Data Analysis (EDA)

EDA was conducted to identify trends and relationships between GDP and revenue for each business. Line plots and scatter plots demonstrated positive relationships between GDP and revenue, with variations across different businesses.

# 4. Model Selection and Implementation

A time series ARIMA model was chosen due to its flexibility in handling time series data and incorporating external regressors such as GDP. Model parameters for each business were determined and evaluated using training data.

# 5. Model Performance

The following table shows evaluation metrics for the models fitted to each business:

|  |  |  |  |
| --- | --- | --- | --- |
| Business | RMSE | MAE | MAPE |
| biz A | 56.10481149293175 | 34.42591431496226 | 0.17549521812634283 |
| biz B | 93.51147484520102 | 64.3253994269465 | 0.1920694277015205 |
| biz C | 70.43695016497477 | 52.64752629910822 | 0.21610603727896108 |
| biz D | 47.13156078800822 | 28.003094164152884 | 0.12436695875594435 |
| biz E | 15.349092463332337 | 13.051076353371357 | 0.13854550600993856 |

# 6. Visualizations of Forecasts

Visualizations comparing historical and forecasted revenue were generated for each business, displaying trends and future projections.

# 7. Fitted vs Actual Values

The following tables show the comparison of fitted and actual revenue values for the last 12 time points for each business:

Fitted vs Actual Values for biz A

|  |  |  |
| --- | --- | --- |
| Date | Actual Revenue | Fitted Revenue |
| 2022-01-31 00:00:00 | 148.8663 | 173.0934704550243 |
| 2022-02-28 00:00:00 | 159.0369 | 173.84342897213557 |
| 2022-03-31 00:00:00 | 384.7892 | 174.59338748924685 |
| 2022-04-30 00:00:00 | 154.8638 | 175.34513162187505 |
| 2022-05-31 00:00:00 | 150.5354 | 176.09509013898634 |
| 2022-06-30 00:00:00 | 149.9374 | 176.84504865609762 |
| 2022-07-31 00:00:00 | 165.4324 | 177.5950071732089 |
| 2022-08-31 00:00:00 | 159.6715 | 178.3449656903202 |
| 2022-09-30 00:00:00 | 154.9401 | 179.09492420743152 |
| 2022-10-31 00:00:00 | 320.6373 | 179.84666834005972 |
| 2022-11-30 00:00:00 | 152.663 | 180.596626857171 |
| 2022-12-31 00:00:00 | 171.6574 | 181.34658537428228 |

Fitted vs Actual Values for biz B

|  |  |  |
| --- | --- | --- |
| Date | Actual Revenue | Fitted Revenue |
| 2022-01-31 00:00:00 | 260.4264 | 287.03196500024114 |
| 2022-02-28 00:00:00 | 253.4472 | 288.22754531369304 |
| 2022-03-31 00:00:00 | 452.6654 | 289.4231256271952 |
| 2022-04-30 00:00:00 | 416.8396 | 290.6215525604852 |
| 2022-05-31 00:00:00 | 263.9718 | 291.81713287398196 |
| 2022-06-30 00:00:00 | 254.1216 | 293.01271318747865 |
| 2022-07-31 00:00:00 | 245.3885 | 294.20829350097534 |
| 2022-08-31 00:00:00 | 256.6187 | 295.40387381447204 |
| 2022-09-30 00:00:00 | 252.9581 | 296.5994541279687 |
| 2022-10-31 00:00:00 | 227.6459 | 297.79788106125943 |
| 2022-11-30 00:00:00 | 258.9309 | 298.9934613747561 |
| 2022-12-31 00:00:00 | 291.0097 | 300.1890416882528 |

Fitted vs Actual Values for biz C

|  |  |  |
| --- | --- | --- |
| Date | Actual Revenue | Fitted Revenue |
| 2022-01-31 00:00:00 | 211.4104 | 231.88329012654 |
| 2022-02-28 00:00:00 | 173.4669 | 232.85757457580524 |
| 2022-03-31 00:00:00 | 419.1362 | 233.83185902507046 |
| 2022-04-30 00:00:00 | 206.4472 | 234.80846319921488 |
| 2022-05-31 00:00:00 | 205.7433 | 235.7827476484801 |
| 2022-06-30 00:00:00 | 196.0996 | 236.75703209774534 |
| 2022-07-31 00:00:00 | 188.7902 | 237.73131654701055 |
| 2022-08-31 00:00:00 | 209.9623 | 238.70560099627576 |
| 2022-09-30 00:00:00 | 198.1683 | 239.67988544554103 |
| 2022-10-31 00:00:00 | 227.191 | 240.65648961968546 |
| 2022-11-30 00:00:00 | 421.1318 | 241.63077406895067 |
| 2022-12-31 00:00:00 | 199.0861 | 242.60505851821588 |

Fitted vs Actual Values for biz D

|  |  |  |
| --- | --- | --- |
| Date | Actual Revenue | Fitted Revenue |
| 2022-01-31 00:00:00 | 187.2114 | 197.37382286050064 |
| 2022-02-28 00:00:00 | 192.0866 | 198.13460984222567 |
| 2022-03-31 00:00:00 | 195.0769 | 198.89539682406354 |
| 2022-04-30 00:00:00 | 186.0012 | 199.65799520346258 |
| 2022-05-31 00:00:00 | 182.4197 | 200.4187821852877 |
| 2022-06-30 00:00:00 | 201.4386 | 201.17956916711262 |
| 2022-07-31 00:00:00 | 175.3467 | 201.94035614893755 |
| 2022-08-31 00:00:00 | 175.5861 | 202.70114313076246 |
| 2022-09-30 00:00:00 | 186.5256 | 203.46193011258742 |
| 2022-10-31 00:00:00 | 183.2711 | 204.2245284919881 |
| 2022-11-30 00:00:00 | 185.2346 | 204.98531547381305 |
| 2022-12-31 00:00:00 | 167.1894 | 205.74610245563798 |

Fitted vs Actual Values for biz E

|  |  |  |
| --- | --- | --- |
| Date | Actual Revenue | Fitted Revenue |
| 2022-01-31 00:00:00 | 110.9473 | 112.51682864379617 |
| 2022-02-28 00:00:00 | 106.3165 | 113.0405304102171 |
| 2022-03-31 00:00:00 | 108.6332 | 113.56423217663803 |
| 2022-04-30 00:00:00 | 103.8085 | 114.08918085202662 |
| 2022-05-31 00:00:00 | 105.3727 | 114.61288261844753 |
| 2022-06-30 00:00:00 | 99.948 | 115.13658438486846 |
| 2022-07-31 00:00:00 | 106.8741 | 115.66028615128938 |
| 2022-08-31 00:00:00 | 102.8978 | 116.1839879177103 |
| 2022-09-30 00:00:00 | 87.8932 | 116.70768968413125 |
| 2022-10-31 00:00:00 | 90.4147 | 117.23263835951984 |
| 2022-11-30 00:00:00 | 106.3109 | 117.75634012594075 |
| 2022-12-31 00:00:00 | 101.038 | 118.28004189236168 |

# 8. Conclusion and Recommendations

The time series regression models for each business demonstrated varying levels of forecast accuracy. The use of GDP as an external variable provided insights into its impact on revenue trends. It is recommended to continue refining these models by exploring additional external variables, adjusting parameters, or using ensemble forecasting methods for better predictive performance.

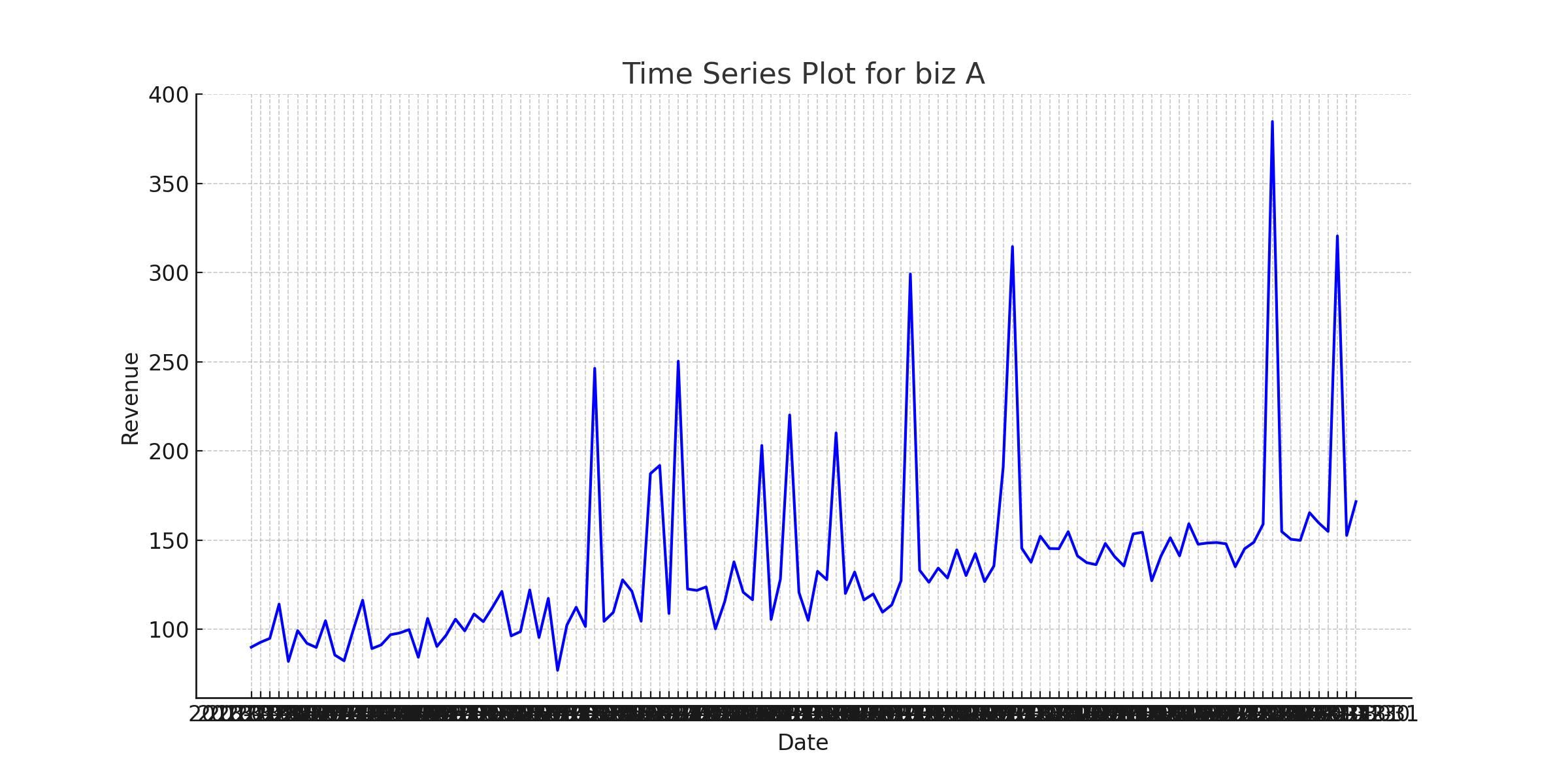


Figure: Biz a time series

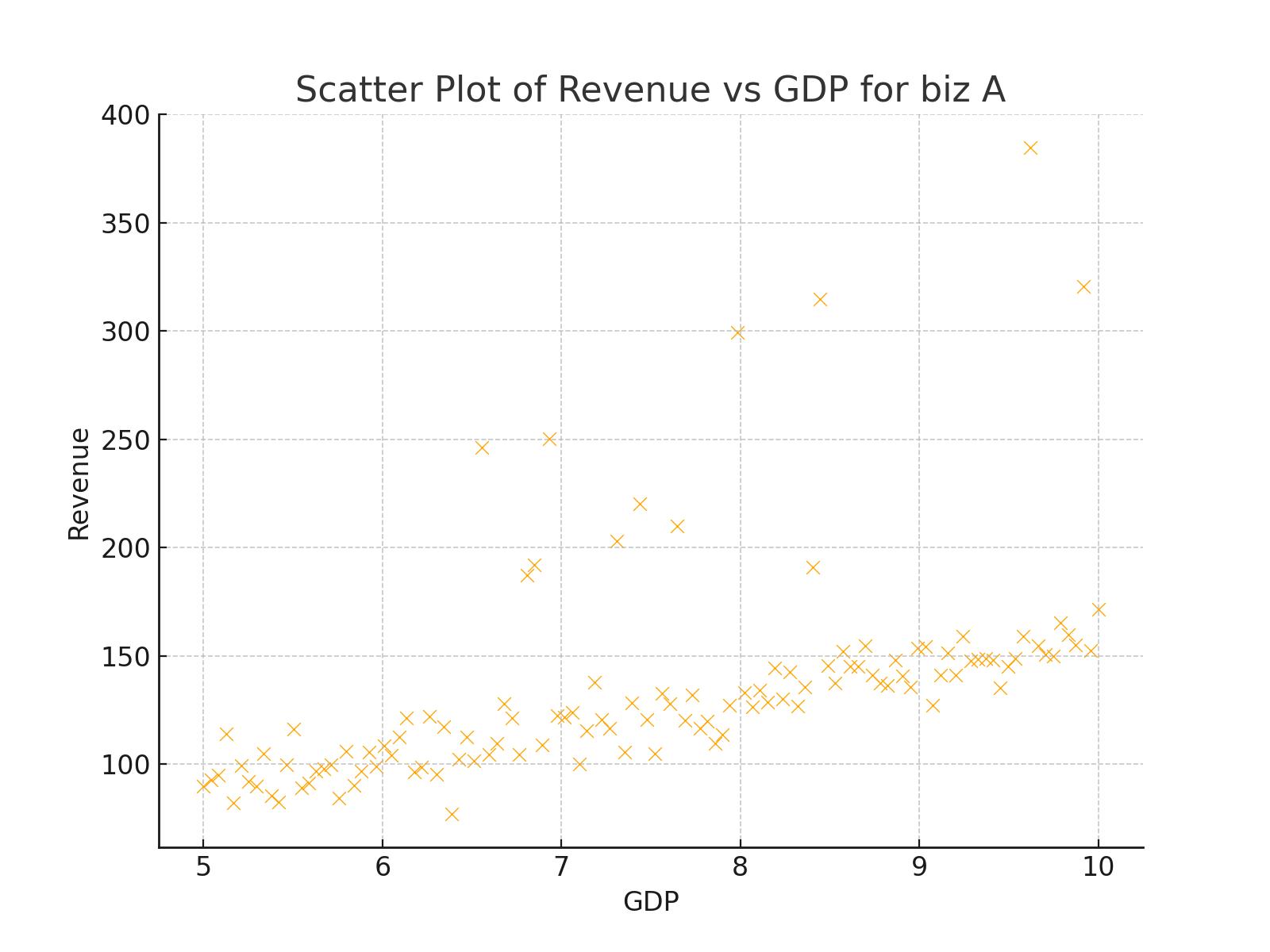


Figure: Biz a scatter

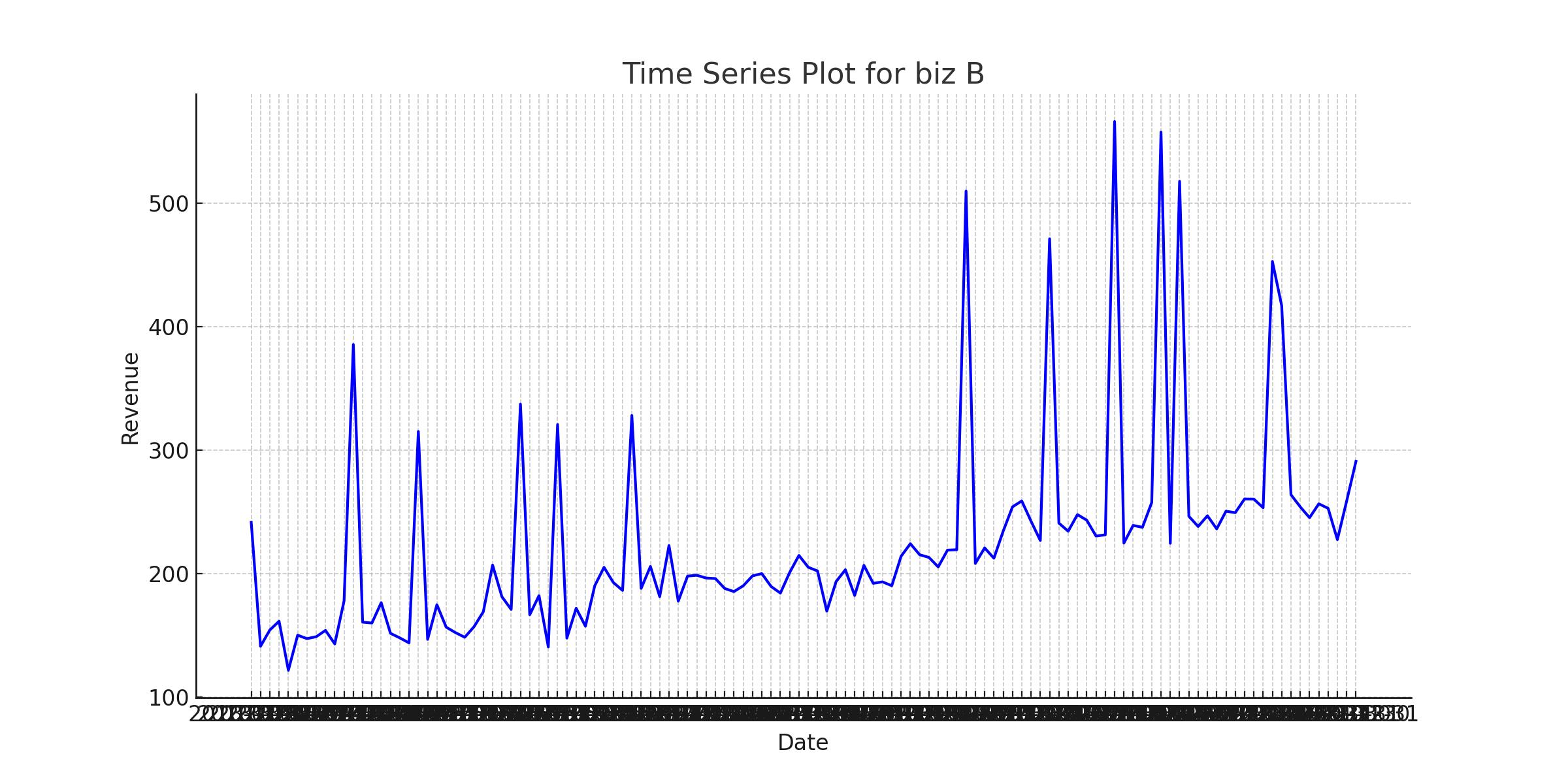


Figure: Biz b time series

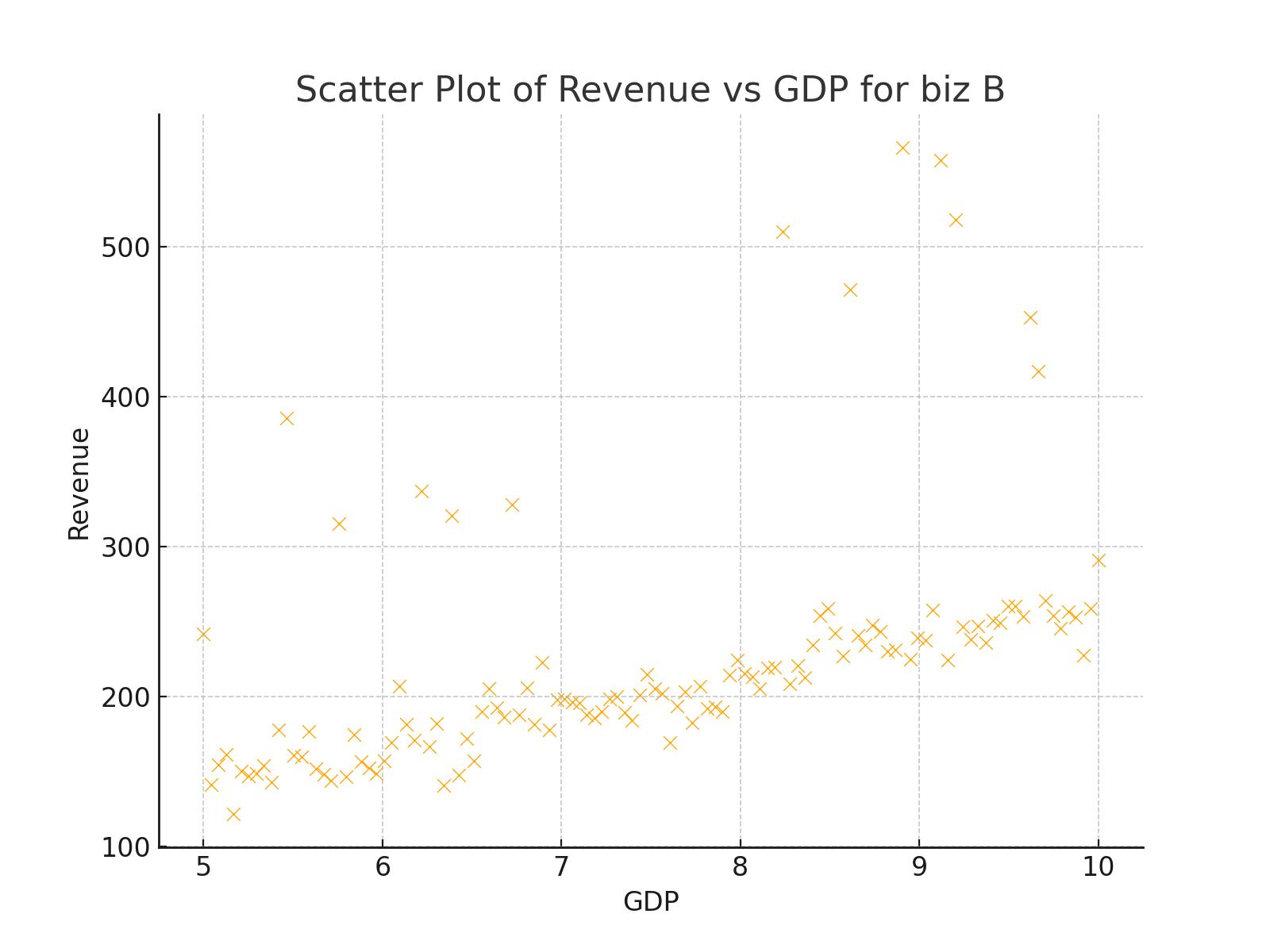


Figure: Biz b scatter

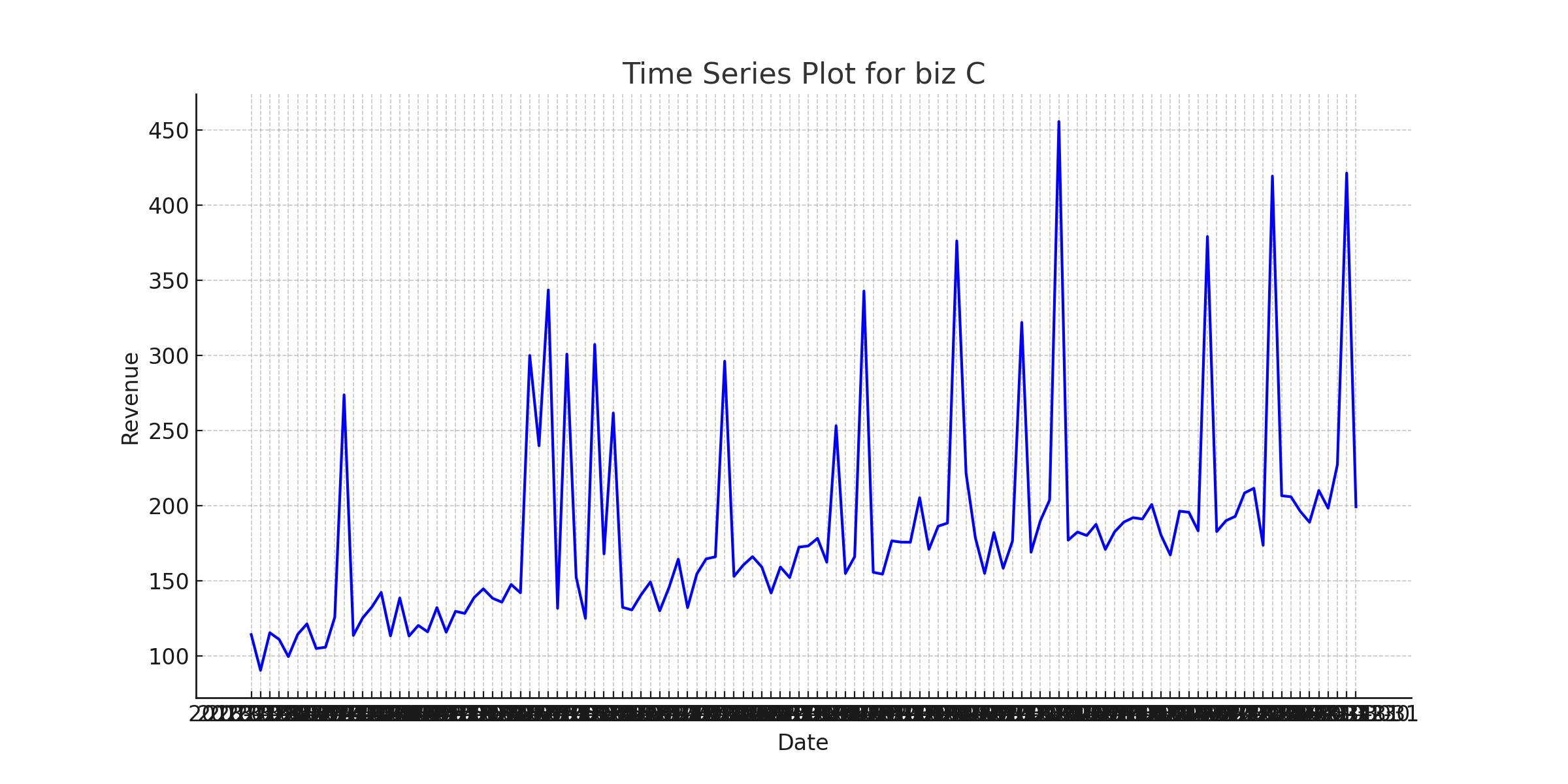


Figure: Biz c time series

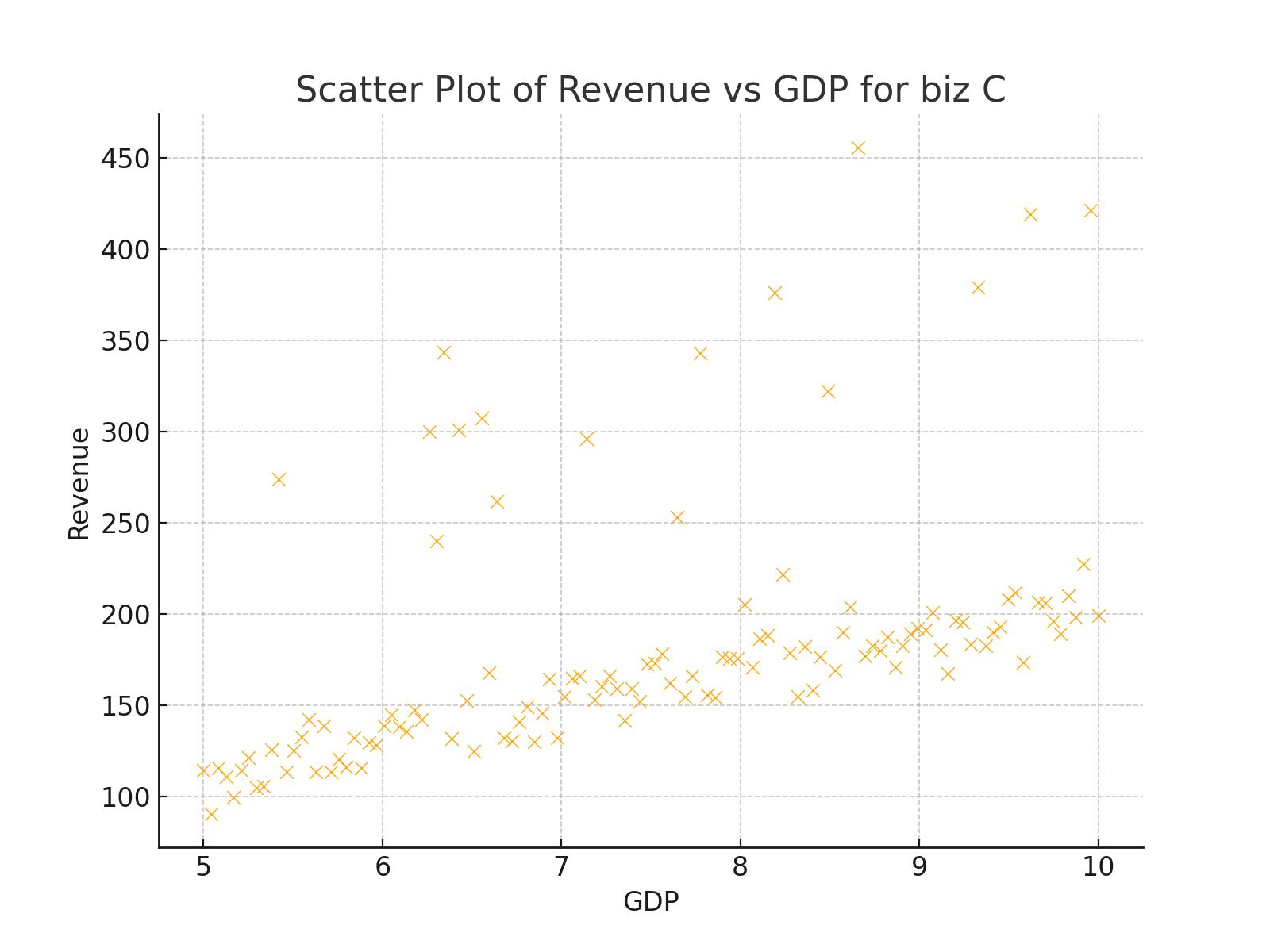


Figure: Biz c scatter

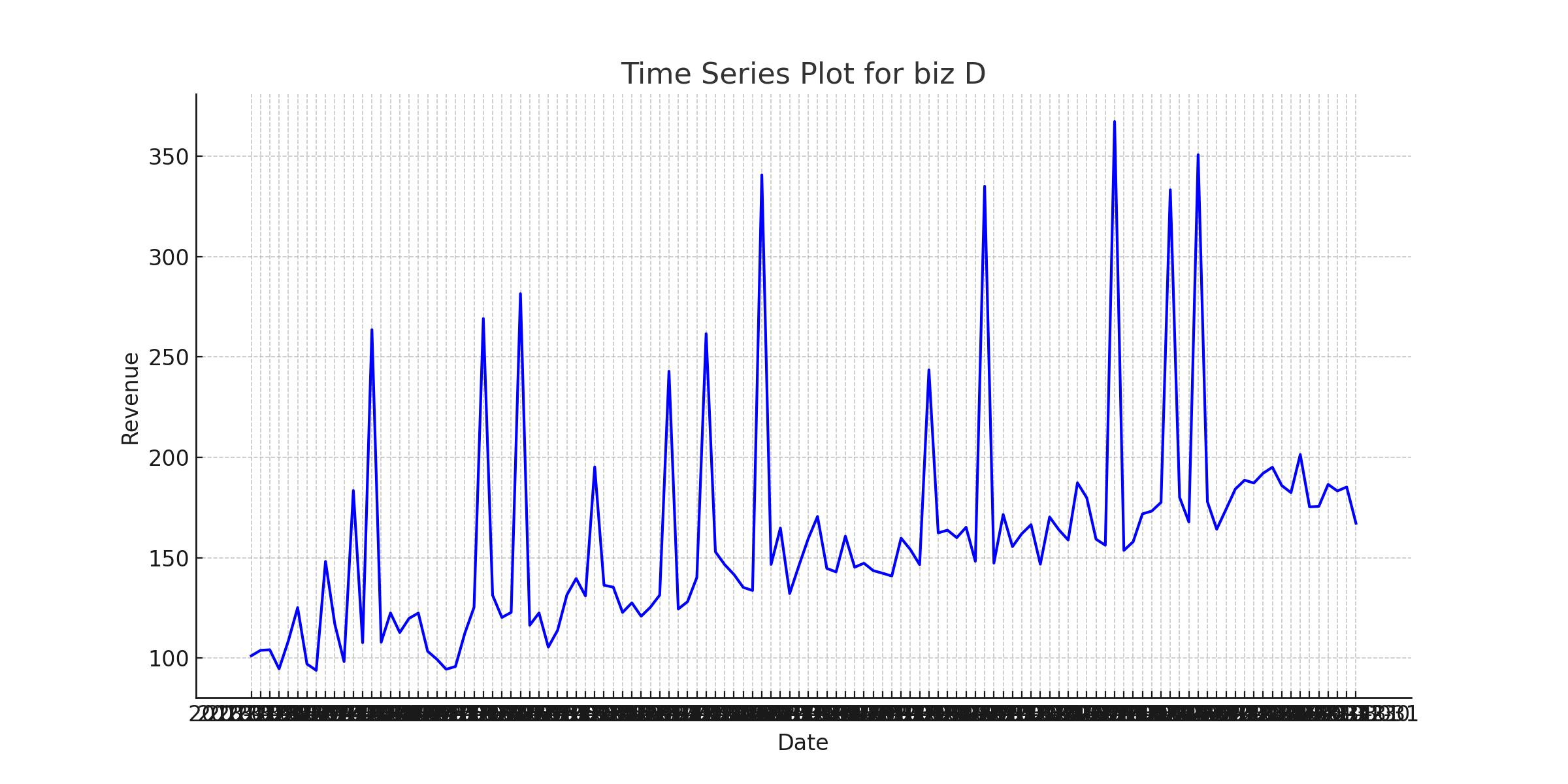


Figure: Biz d time series

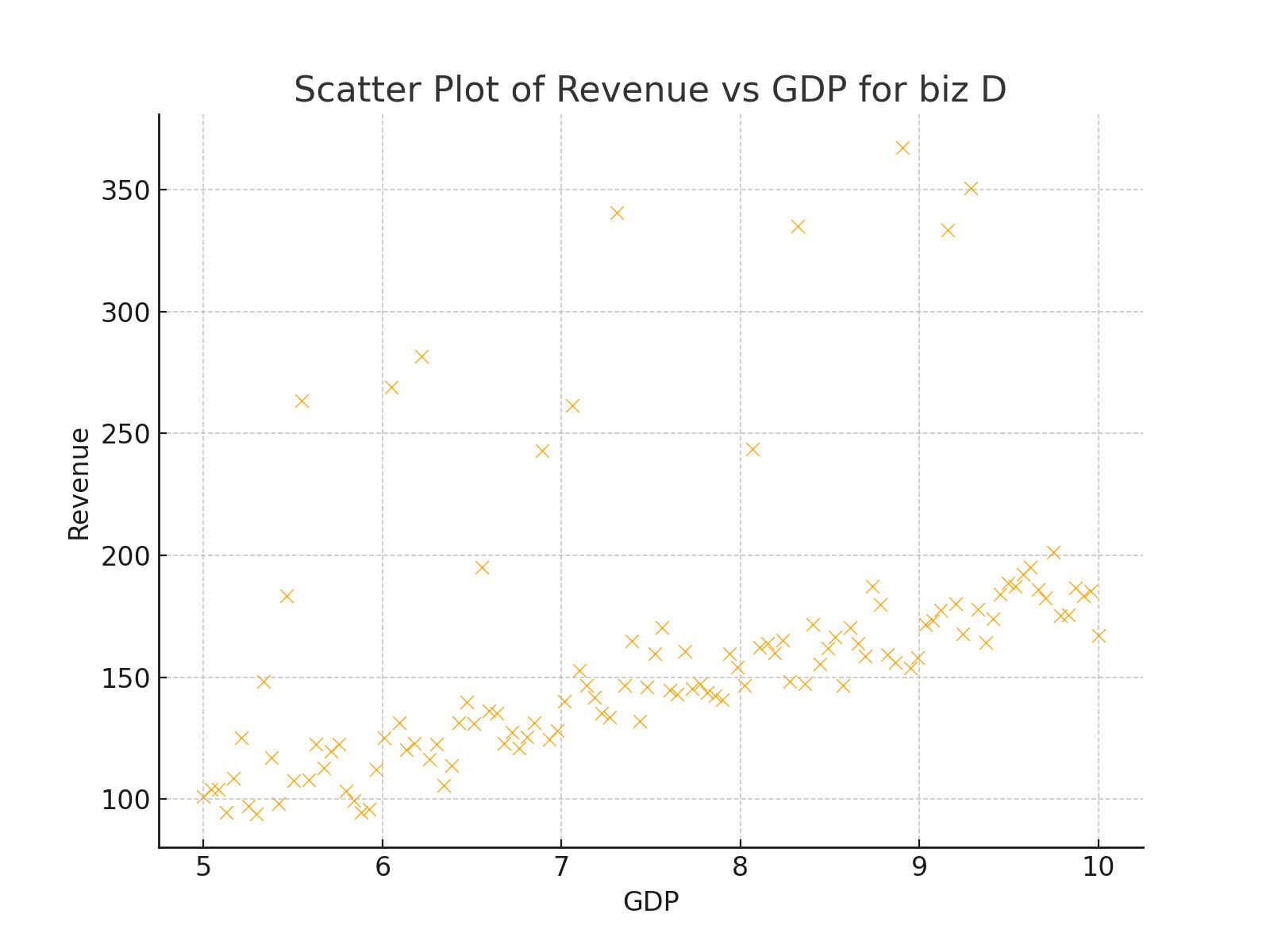


Figure: Biz d scatter

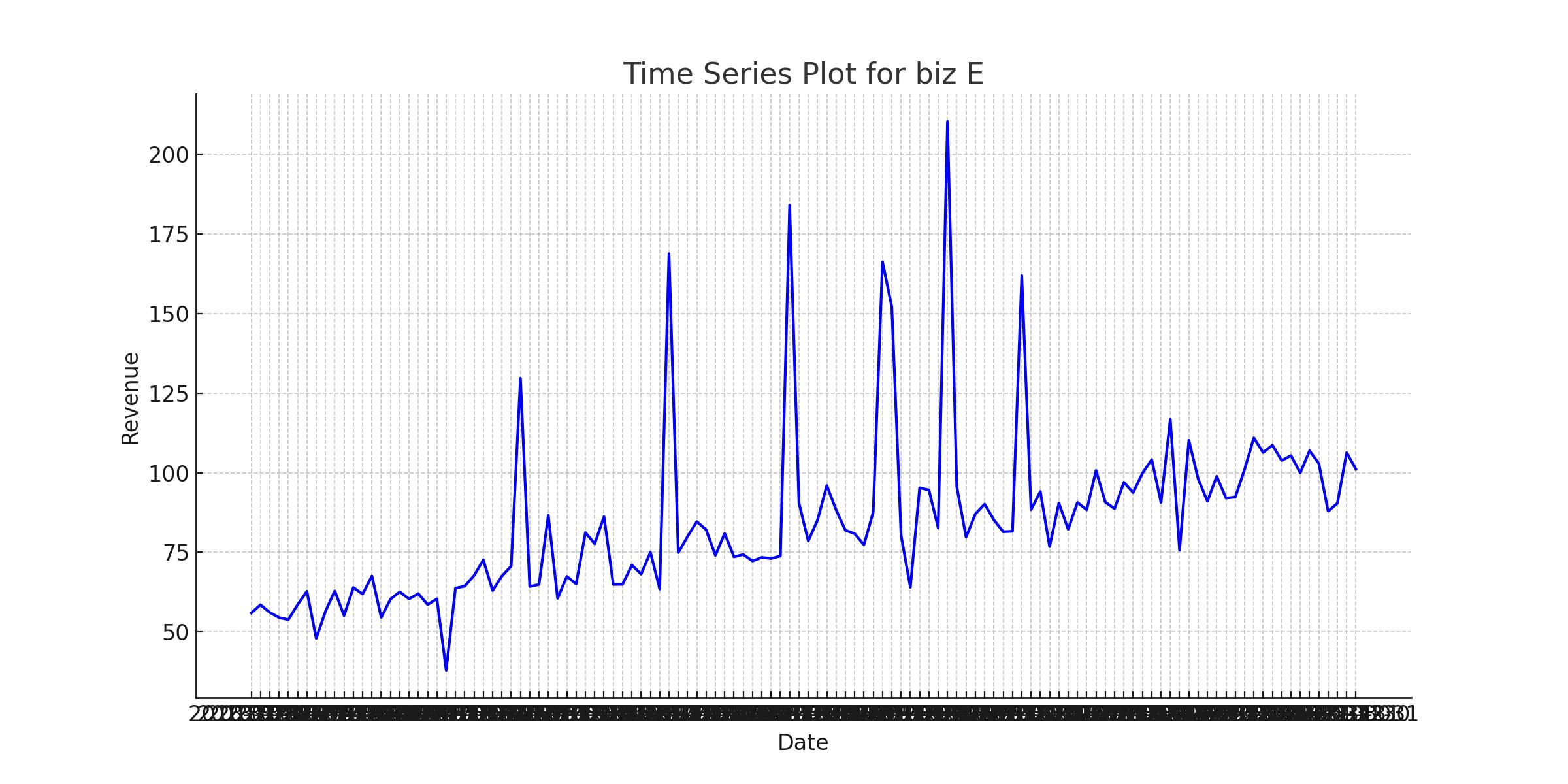


Figure: Biz e time series

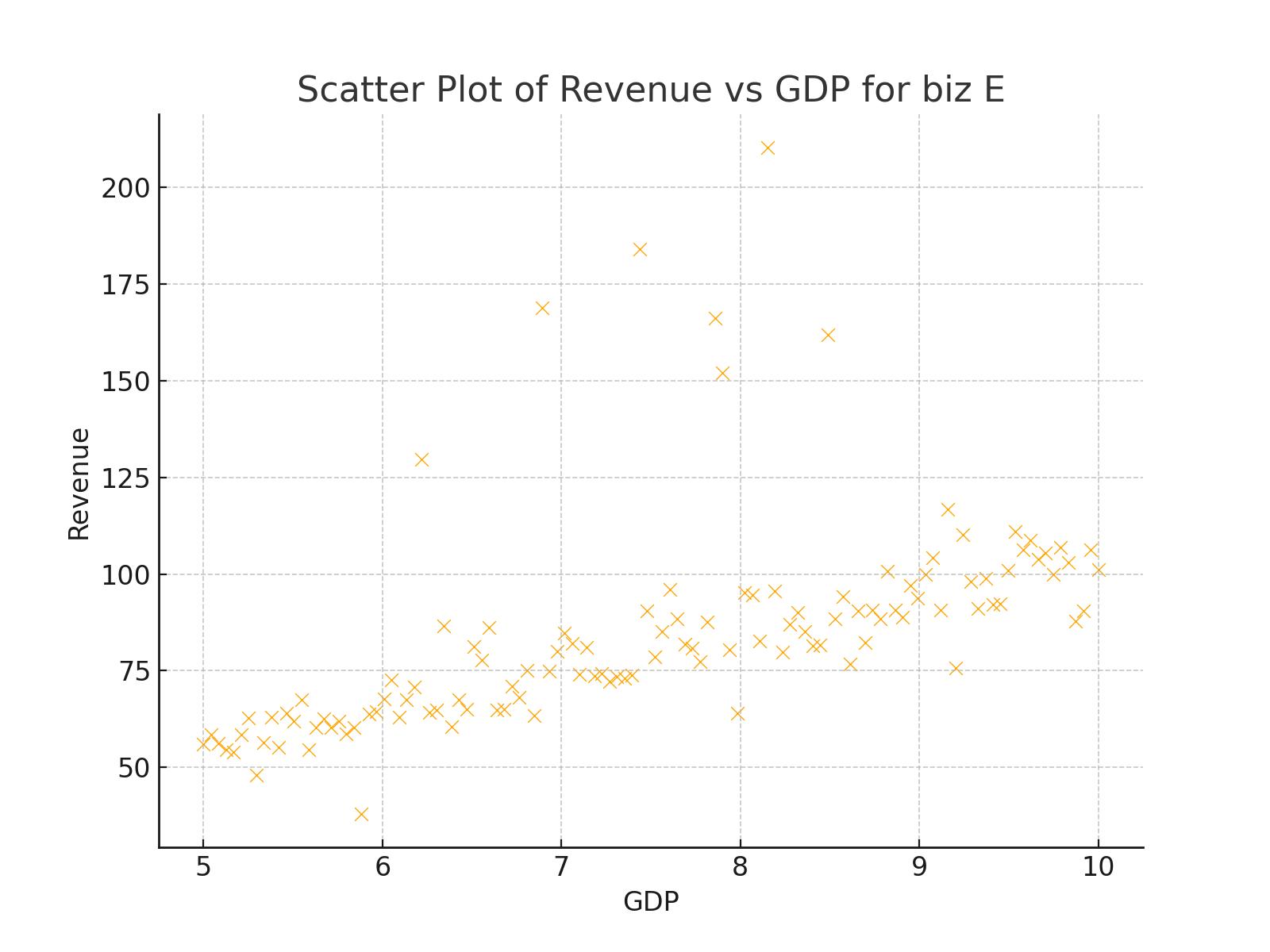


Figure: Biz e scatter