Future Predictions for Top 10 Obligation Products

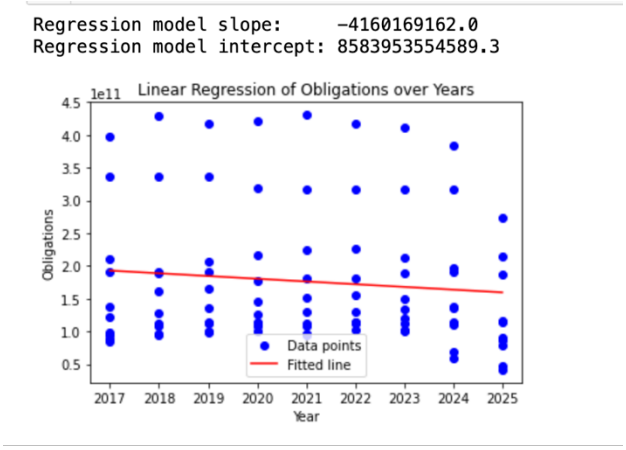
This section makes use of data from USASpending.gov for federal spending by product and service 2017-2025, the first ten obligations by product and service for each year, and data from Word Bank Group, including GDP growth (annual%), military spending (% of GDP), and R&D expenditure (% of GDP), 1960-2025. Military spending (% of GDP) and R&D expenditure (% of GDP) from 1960 to 2025.

Data processing

Obligations: Delete the dollar sign and turn it into a number

US data：Delete data from other countries and fill null with mean

Linear Regression



Regression Model Slope: -4160169162.0

This means that the estimated value of the debt reduces by approximately 4.16 trillion (billion) units with each consecutive year, exhibiting a downward tendency.

Regression Model Intercept: 8583953554589.3

This is the expected value of the obligation at year 0 (the theoretical starting point), which is around 8.58 MU.

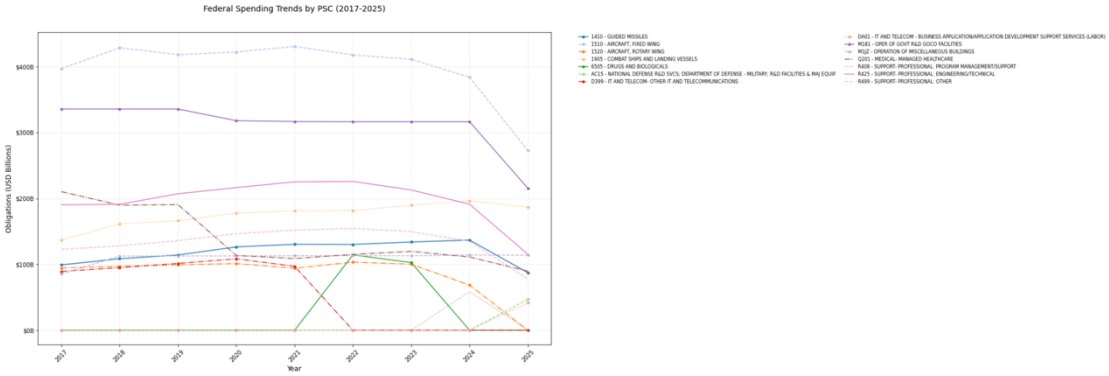
Regression line: The red fitted line is the outcome of a linear regression of these data, demonstrating the expected trend of responsibilities over time.

Trend: As seen in the graph, the slope of the regression line is negative, indicating that

the obligation data is likely to decrease as the number of years grows.

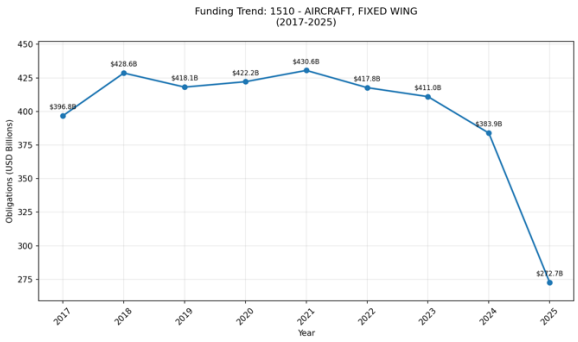
Analysing historical data

Using historical data, the first ten product responsibilities for each year from 2017 to 2025 were plotted to provide a better picture of the general trends for these products.



The data for each of these items was then plotted to provide a better understanding of their respective situations.

For example, the obligation trend of "1510 - AIRCRAFT, FIXED WING" has consistently been very high.



Find out most frequent keywords:

```
Keyword: AIRCRAFT,, Products: ['1510 - AIRCRAFT, FIXED WING', '1520 - AIRCRAFT, ROTARY WIN
G']
Keyword: SUPPORT-, Products: ['R425 - SUPPORT- PROFESSIONAL: ENGINEERING/TECHNICAL', 'R499
- SUPPORT- PROFESSIONAL: OTHER', 'R408 - SUPPORT- PROFESSIONAL: PROGRAM MANAGEMENT/SUPPOR
T']
Keyword: IT, Products: ['D399 - IT AND TELECOM- OTHER IT AND TELECOMMUNICATIONS', 'DA01 - I
T AND TELECOM - BUSINESS APPLICATION/APPLICATION DEVELOPMENT SUPPORT SERVICES (LABOR)']
```

Based on the above data, the following analyses can be derived:

The aviation industry is critical for the United States. The category "AIRCRAFT" occurs twice in the top ten budget responsibilities, emphasizing its importance. From 2017 to 2025, aviation has regularly ranked first among budgetary responsibilities, demonstrating the United States' strong priority for this sector and its critical role in national defense and the economy.

According to a review of IT products, the United States is transitioning from an infrastructure-driven to an applications and innovation-focused approach. The product D399, which symbolizes IT AND TELECOMMUNICATIONS, was among the top 10 obligations only lasts until 2021. In contrast, DA01, which stands for BUSINESS APPLICATION/APPLICATION DEVELOPMENT, first appeared in the top ten rankings in 2025. D399 primarily provides telecommunications services such as technical assistance and network services, whereas DA01 focuses on business application support services. This trend underscores the fact that businesses are increasingly prioritizing digital transformation, with a significant emphasis on application development and technical support to improve efficiency and competition.

KNN model

```
Accuracy (per fold)=  [0.61111111 0.66666667 0.38888889 0.5        0.55555556]
Accuracy (average)=  0.544 ( 0.096 )

Confusion Matrix:
================
TN= 26 FP= 23 FN= 18 TP= 23
Recall/Sensitivity=  0.561
Specificity=  0.531
Precision=  0.5
Accuracy (per fold, PCA)=  [0.61111111 0.66666667 0.38888889 0.5        0.55555556]
Accuracy (average, PCA)=  0.544 ( 0.096 )

Confusion Matrix (PCA):
================
TN= 26 FP= 23 FN= 18 TP= 23
Recall/Sensitivity=  0.561
Specificity=  0.531
Precision=  0.5
```

The accuracy of each fold of the KNN model is relatively low, with an average accuracy of just 0.544 and a high standard deviation. This indicates that the data is not satisfactory when analyzing the product data in isolation.

Logistic model

A new US_data csv with GDP growth (annual %), military spending (% of GDP) and R&D expenditure (% of GDP) from 1960 to 2024 has been added to improve the modelling accuracy.

Feature engineering：

Due to the large number of defence products in the first 10 data and the presence of military spending in the US data, an is_defense_related method is added as a classification criterion to help the analysis.

```
Accuracy: 0.81
ROC-AUC: 0.78

2025 Product predict result：
    Year                                              Product  %_of_Increase
77  2025                  COMBAT SHIPS AND LANDING VESSELS       0.214177
78  2025                                   GUIDED MISSILES       0.201865
84  2025       SUPPORT- PROFESSIONAL: ENGINEERING/TECHNICAL       0.102936
83  2025               OPERATION OF MISCELLANEOUS BUILDINGS       0.099467
85  2025                    SUPPORT- PROFESSIONAL: OTHER       0.074721
81  2025  NATIONAL DEFENSE R&D SVCS; DEPARTMENT OF DEFEN...       0.067595
79  2025  IT AND TELECOM - BUSINESS APPLICATION/APPLICAT...       0.063773
76  2025                              AIRCRAFT, FIXED WING       0.059050
80  2025                   MEDICAL- MANAGED HEALTHCARE       0.050942
82  2025                  OPER OF GOVT R&D GOCO FACILITIES       0.020533
```

The accuracy of the model was 81%, ROC-AUC, 0.78, indicating that the model has good predictive power.

‘%_of_Incresae’ section shows the different products predicted by the model and their probability of increasing in the future.

Because of the declining trend in overall Obligation expenditure, as well as the relatively low GDP and Military spending (% of GDP) data, future spending increases for each Obligation product are typically unlikely. This means that, despite relatively high anticipated growth rates for particular goods such as "COMBAT SHIPS AND LANDING VESSELS" and "GUIDED MISSILES", the general market situation continues to exert pressure on spending growth.

Summary

This analysis examines the top ten obligated products in the United States from 2017 to 2025, with a focus on federal spending, GDP growth, military spending, and research and development investment. A linear regression model indicates a decreased trend in obligation spending. Analyses based on historical data show the importance of the aviation industry to the US. Furthermore, an analysis of IT goods

reveals that the market is changing from telecommunications services to business application development, reflecting the trend of digital transformation. To improve model accuracy, new US data is included for feature engineering, and the logistic regression model's accuracy is increased to 0.81, demonstrating that the model has high predictive power.