HAP/AIT-622

- *PROF.BRETT BERLIN*

FINAL PROJECT
PART-2
DREAM TEAM

TEAM MEMBERS

CAMERON HURLEY (G00983900)
DESHIK SAI KAKARLA (G1324749)
MOUNIKA JAGGAMPUDI(G01359781)
NIHAR KALUVALA (G01353048):

Table of Contents

A) Assume you are responding to the collecting organization's RFP (Request for Proposal)

We will do a detailed analysis of the Medicare Part D Spending by prescription dataset to study the trends and patterns of prescription spending for Medicare Part D drugs. This study will help us to identify the biggest-spending manufacturers, commonly prescribed drugs in Part D, and the drugs with the higher spending increase rate over a period.

To understand the dataset better, we will use a variety of statistical analyses and visualizations. Our objective is to identify significant differences in drug spending across beneficiary categories using descriptive statistics and inferential statistics. Additionally, various plots such as histograms, scatterplots, and box plots will be used to identify the outliers and trends in drug spending.

The aim of this analysis is to provide insights which aid to improve patient outcomes, make better policy decisions which in turn improve the efficiency of healthcare systems. The results will also help in reconsidering the pricing policies of various medicinal products which promote affordability and accessibility. This also helps government in identifying ways to reduce spending on drugs without compromising the quality. Finally, our analysis intends to improve the efficiency and equity of healthcare so that all beneficiaries can receive high-quality care.

B) The following list of stakeholders involved in the project can be identified:

Beneficiaries of Medicare:

The Medicare Part D capital structure straightforwardly impacts these individuals' access to formula drug coverage. Medicare Part D coverage is available to in addition 47% of Medicare benefits, in accordance with the Centers for Medicare & Medicaid Services (CMS).

Pharmaceutical companies:

These businesses directly benefit from Medicare Part D's spending practices because the program includes prescription drugs made by them. Medicare Part D accounted for about 14% of US prescription drug spending in 2018, according to the Kaiser Family Foundation.

Health care providers:

The expenditure patterns of the program may have an impact on these medical professionals because they are the ones who prescribe the pharmaceuticals covered by Medicare Part D. In 2020, there were roughly 1.3 million doctors and other healthcare professionals registered in Medicare, according to the CMS.

Insurance companies:

These are the corporations that handle the administration of Medicare Part D programs and bargain rates with drug manufacturers, and they could be impacted by the program's expenditure trends. The CMS reports that 221 businesses offered 914 Medicare Part D plan contracts in 2021.

Government Agencies:

These organizations are in charge of managing Medicare Part D and ensuring that the scheme runs smoothly and successfully. The program's expenditure patterns might potentially have an effect on them. The Centers for Medicare and Medicaid Services, a division of the U.S. health department, manages Medicare Part D.

Taxpayers:

In 2020, $651 billion in administration services was used to finance Medicare according to the CMS. They have a financial stake in keeping Medicare running as efficiently as possible because they pay taxes to fund it.

Researchers and policy analyst:

These are the people and organizations that might evaluate the efficacy of the program and suggest policy changes using the information on Medicare Part D spending. For instance, the Kaiser Family Foundation frequently releases papers and analysis on the trends and patterns of Medicare Part D spending.

C) Are there any privacy, quality, or other issues with this data?

The amount that Medicare spent on prescription pharmaceuticals under the Part D program is detailed in the Medicare Part D Spending by Drug dataset. There may be privacy and quality concerns, just as there are with any dataset.

Privacy:

The dataset includes potentially delicate and private personal health information. If this information is made public, there can be secrecy violations and negative consequences for the parties troubled. The data has happened de-identified to look after the secrecy of the subjects.

Quality:

Errors in recording or deal with could have an affect the data's exactness. It is important to consider the data's fullness, constancy, and exactness before using it for reasoning or in charge. The dataset ability not be representative of the complete population on account of choice bias, in the way that dissimilarities in the characteristics of cases contained in the data distinguished to those not

included. It's possible that the data, that only covers the year 2018, doesn't correctly indicate the trends in Medicare giving immediately.

D) Requirements:

i)

People: You will need a company of very qualified professionals accompanying experience in data learning, data architecture, data analysis, and data image. You might also need field specialists, analysts for business, and project leaders, established the opportunity and degree of trouble of your project.

Technology: To handle and analyze enormous quantities of data, you will require a variety of hardware and software equipment. This includes:

Hardware: To cope with the processing requirements of big data, you must have a cluster of outstanding performance computers. Servers, storage units, and networking hardware may all be included under this category.

Software: For different duties, such as data integration, processing, analysis, and visualization, you will require an array of software resources. Hadoop, Spark, Hive, Pig, and NoSQL databases are now just a few of the technologies that are commonly employed for big data projects.

You need a detailed project plan, an explicit set of targets and objectives, and an effective data management approach in along with these resources. While organizing your big data project, you ought to also consider data security, scalability, and reliability into consideration.

ii)

There are various determinants to take into concept when determining when to build or buy supplies for a large data project. Among some of ultimate main one are filed beneath:

Cost: It might be costly to develop a big data solution from the bottom up, both economically and regarding time. In contrast, buying a solution may be more affordable.

Time to market: Buying a solution might be quicker for deployment than creating one from the beginning.

Customizability: You have greater influence over a product's characteristics and features since you create it from starting up. On the contrary, buying a solution may render it harder to modify it to suit your specific needs.

Support and maintenance costs: Firms who lack the ability to handle their own software might find it advantageous to buy an approach which offers reinforce and servicing.

Below are some ideas for when to develop or buying a big data solution based around these considerations:

Build if: Off-the-shelf items are not able to meet the specific requirements of your company's requirements.

You have the opportunity, money, and talent necessary to create and sustain an approach.

You need to have complete influence over the features and abilities of your goods.

Buy if: the organization requires an immediate resolution but possesses few assets.

There are several already made choices readily accessible that will meet your needs.

You choose to depend on a supplier for assistance and servicing.

E) Describe the relevant metadata (types of data in the dataset)

The metadata in this dataset consist of a variety of different data types. The most consistent data type here are numerical data types that are either integer, float, or decimal types. The second most appeared data type would be the Boolean data type but instead of the traditional "TRUE" or "FALSE" values they are "1" and "0" values. The last data type that is present in this dataset are string values. These values here are used to describe the drug name and their respective manufacturers.

F)the appropriate types of statistical analyses that would be performed.

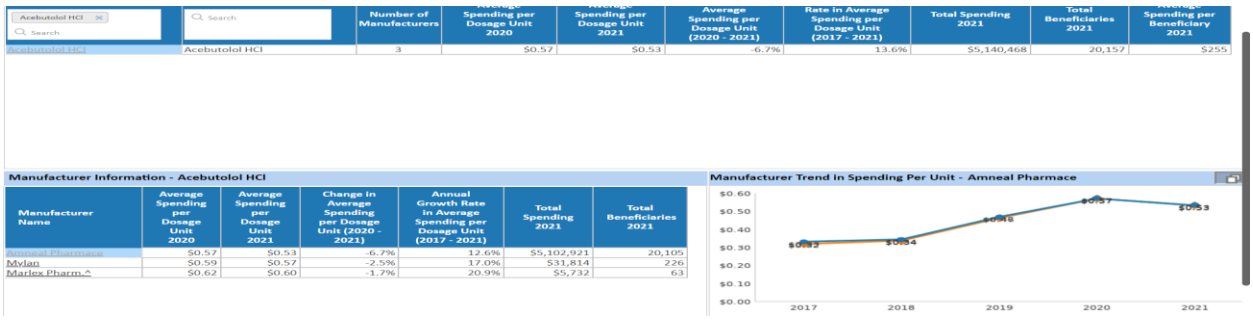
To explore and understand the patterns and trends in the dataset and to test the hypothesis to make valid inferences, statistical analyses play a foundation role in building models and making efficient predictions. Correlation analysis gives insight to which variables like the different vaccines and companies have strong relations. Also, Descriptive analysis like mean, and standard deviation of the dosage units in each of the spending patterns of Part D drugs. Furthermore, regression analysis is used to explore any relative importance of various predictors.

G) the appropriate types of visualizations that would be performed

As visualizations give clear and an easy way of deeper insights to the trends and patterns of a dataset, for identifying the spending patterns of these drugs, bar charts, scatterplots, box plots can be used to illustrate and give efficient analysis through these visualizations as the dataset contains various units of dosages and its variety of manufacturers. Line charts and dashboards can also be used to study about the spending patterns of individual molecules by different manufacturers some of the examples have been showed below.

Dashboard which shows manufacturer spending per unit for different drugs

Spending per unit by Amneal Pharmace for Acebutolol HCL



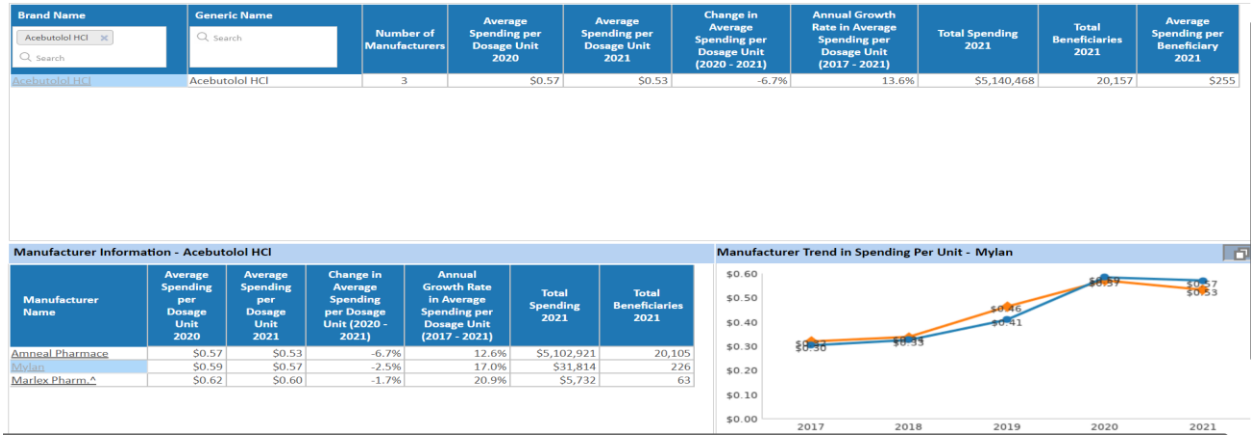
Brand Name	Generic Name	Number of Manufacturers	Average Spending per Dosage Unit 2020	Average Spending per Dosage Unit 2021	Change in Average Spending per Dosage Unit (2020 - 2021)	Annual Growth Rate in Average Spending per Dosage Unit (2017 - 2021)	Total Spending 2021	Total Beneficiaries 2021	Average Spending per Beneficiary 2021
Acebutolol HCl	Acebutolol HCl	3	$0.57	$0.53	-6.7%	13.6%	$5,140,468	20,157	$255

Manufacturer Information - Acebutolol HCl

Manufacturer Name	Average Spending per Dosage Unit 2020	Average Spending per Dosage Unit 2021	Change in Average Spending per Dosage Unit (2020 - 2021)	Annual Growth Rate in Average Spending per Dosage Unit (2017 - 2021)	Total Spending 2021	Total Beneficiaries 2021
Amneal Pharmace	$0.57	$0.53	-6.7%	12.6%	$5,102,921	20,105
Mylan	$0.59	$0.57	-2.5%	17.0%	$31,814	226
Marlex Pharm.^	$0.62	$0.60	-1.7%	20.9%	$5,732	63

Manufacturer Trend in Spending Per Unit - Amneal Pharmace

Spending per unit by Mylan for Acebutolol HCL



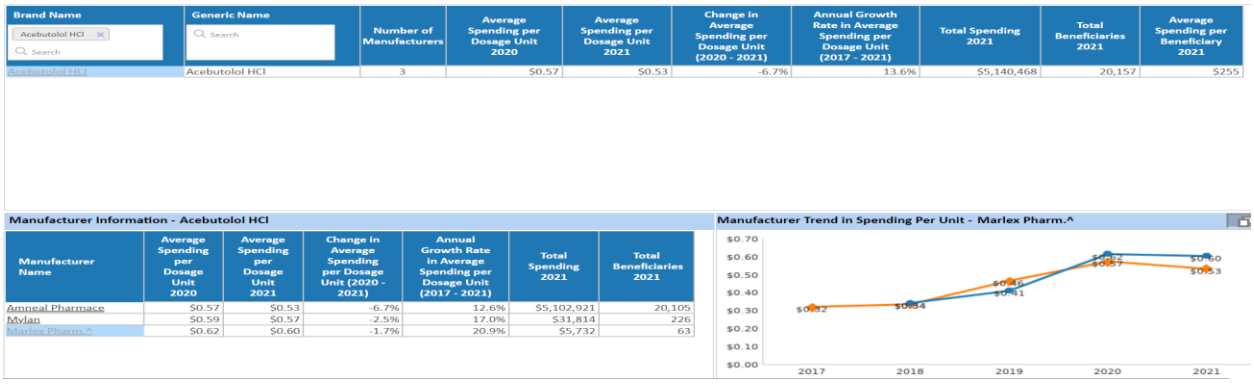
Brand Name	Generic Name	Number of Manufacturers	Average Spending per Dosage Unit 2020	Average Spending per Dosage Unit 2021	Change in Average Spending per Dosage Unit (2020 - 2021)	Annual Growth Rate in Average Spending per Dosage Unit (2017 - 2021)	Total Spending 2021	Total Beneficiaries 2021	Average Spending per Beneficiary 2021
Acebutolol HCl	Acebutolol HCl	3	$0.57	$0.53	-6.7%	13.6%	$5,140,468	20,157	$255

Manufacturer Information - Acebutolol HCl

Manufacturer Name	Average Spending per Dosage Unit 2020	Average Spending per Dosage Unit 2021	Change in Average Spending per Dosage Unit (2020 - 2021)	Annual Growth Rate in Average Spending per Dosage Unit (2017 - 2021)	Total Spending 2021	Total Beneficiaries 2021
Amneal Pharmace	$0.57	$0.53	-6.7%	12.6%	$5,102,921	20,105
Mylan	$0.59	$0.57	-2.5%	17.0%	$31,814	226
Marlex Pharm.^	$0.62	$0.60	-1.7%	20.9%	$5,732	63

Manufacturer Trend in Spending Per Unit - Mylan

Spending per unit by Marlex Pharm for Acebutolol HCL



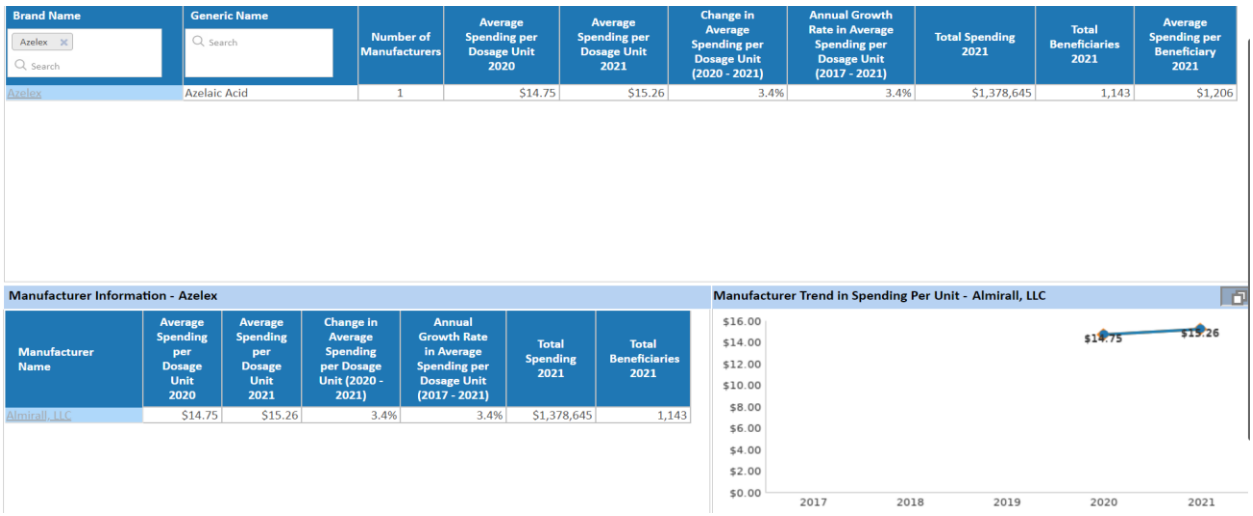
Brand Name	Generic Name	Number of Manufacturers	Average Spending per Dosage Unit 2020	Average Spending per Dosage Unit 2021	Change in Average Spending per Dosage Unit (2020 - 2021)	Annual Growth Rate in Average Spending per Dosage Unit (2017 - 2021)	Total Spending 2021	Total Beneficiaries 2021	Average Spending per Beneficiary 2021
Acebutolol HCl	Acebutolol HCl	3	$0.57	$0.53	-6.7%	13.6%	$5,140,468	20,157	$255

Manufacturer Information - Acebutolol HCl

Manufacturer Name	Average Spending per Dosage Unit 2020	Average Spending per Dosage Unit 2021	Change in Average Spending per Dosage Unit (2020 - 2021)	Annual Growth Rate in Average Spending per Dosage Unit (2017 - 2021)	Total Spending 2021	Total Beneficiaries 2021
Amneal Pharmace	$0.57	$0.53	-6.7%	12.6%	$5,102,921	20,105
Mylan	$0.59	$0.57	-2.5%	17.0%	$31,814	226
Marlex Pharm.^	$0.62	$0.60	-1.7%	20.9%	$5,732	63

Manufacturer Trend in Spending Per Unit - Marlex Pharm.^

Spending per unit by Almirall for Azelex



Brand Name	Generic Name	Number of Manufacturers	Average Spending per Dosage Unit 2020	Average Spending per Dosage Unit 2021	Change in Average Spending per Dosage Unit (2020 - 2021)	Annual Growth Rate in Average Spending per Dosage Unit (2017 - 2021)	Total Spending 2021	Total Beneficiaries 2021	Average Spending per Beneficiary 2021
Azelex	Azelaic Acid	1	$14.75	$15.26	3.4%	3.4%	$1,378,645	1,143	$1,206

Manufacturer Name	Average Spending per Dosage Unit 2020	Average Spending per Dosage Unit 2021	Change in Average Spending per Dosage Unit (2020 - 2021)	Annual Growth Rate in Average Spending per Dosage Unit (2017 - 2021)	Total Spending 2021	Total Beneficiaries 2021
Almirall, LLC	$14.75	$15.26	3.4%	3.4%	$1,378,645	1,143

H) Estimate the time/effort required for the study

We would estimate that this study required months or maybe even a year or two to fully finish. The longest part of this study was most likely the data collection phase as well as the data cleaning stage. This dataset contains data from many different databases, and it probably took quite a while to select what data to pull and merge the datasets once they were queried. With datasets coming from different databases, it most likely took months to look through all the data and perform data cleaning to make everything uniform.

I) the expected value/benefits to the organization for conducting the study

The collecting organization can get significant insights into the factors that drive drug spending in the Medicare program by performing a detailed analysis of the Medicare Part D Spending by Drug dataset. This data can be used to discover cost-cutting opportunities, such as promoting the use of generic medications or negotiating cheaper prices with manufacturers. By studying these spending trends and patterns we can identify the areas that has limited access to certain drugs, which might lead to opportunities for improving policy measures that improve patient outcomes.

The results of this study can be used to take decisions on policies related to drug pricing and access to Medicare. For example, the study could reveal discrepancies in drug use across different populations, which could drive policy targeted at enhancing equity and eliminating inequities. This study will help the organization to implement new pricing policies which are dependent on the utilization of healthcare services by patients and the patient outcomes.

To sum up, this study has many advantages, as it improves patient outcomes by taking better policy decisions. The organization can make accurate decisions that benefit patients and the healthcare system by having better insight on the factors that drive prescription spending in the Medicare program.

REFERENCES

1. Medicare Advantage/Part D Contract and Enrollment Data | CMS. (2023, March 20). Www.cms.gov. https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/mcradvpartdenroldata
2. Centers for Medicare and Medicaid Services. (2019). About CMS | CMS. Cms.gov. https://www.cms.gov/About-CMS/About-CMS
3. Centers for Medicare & Medicaid Services. (2021). Medicare Part D Drug Spending Dashboard. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Information-on-Prescription-Drugs/index
4. National Institutes of Health. (2021). De-identification and Re-identification of Protected Health Information.
5. https://privacyruleandresearch.nih.gov/pr_08.asp
6. Apache Hadoop. (n.d.). Retrieved April 21, 2023, from
7. https://hadoop.apache.org/
8. Google. (n.d.). Dataproc | google cloud. Google. Retrieved April 21, 2023, from https://cloud.google.com/dataproc

DATA SET LINK:

Medicare Part D Spending by Drug. (2023, March 6). Data.gov; Centers for Medicare & Medicaid Services. https://catalog.data.gov/dataset/medicare-part-d-spending-by-drug-401d2

DATA DICTIONARY LINK

https://data.cms.gov/resources/medicare-part-d-spending-by-drug-data-dictionary