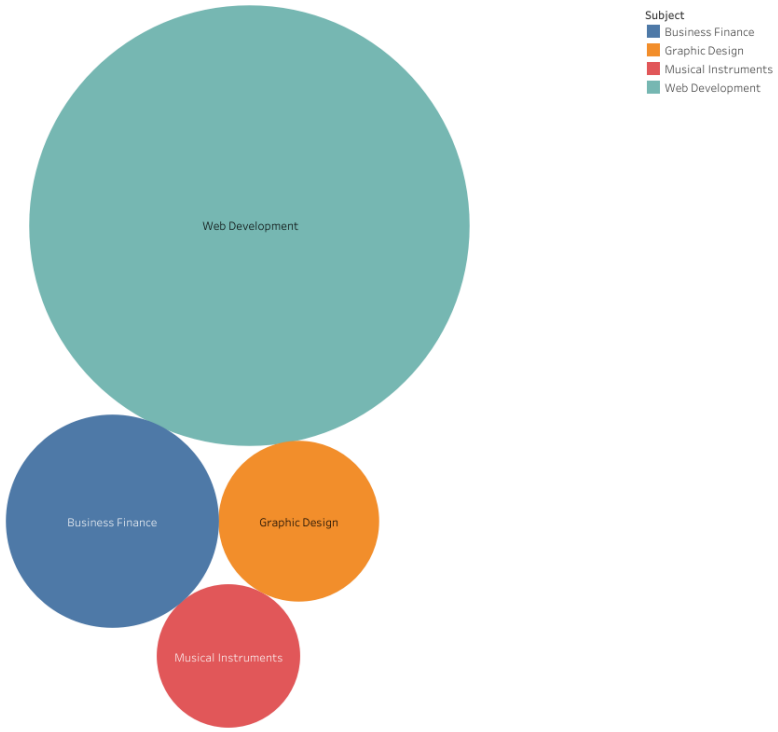


DATA ANALYSIS PORTFOLIO

Professional Background

My name is James Brown, and I'm a Data Entry Analyst looking to flex my talents to identify new growth strategies for your organization as a Junior Data Analyst. I've always been fascinated by numbers, and working in data analytics has been a long-term goal of mine. My career in the food-tech business has taught me to think strategically about problems and identify solutions. I believe this experience has prepared me for Chevron Junior Data Analyst role.

Your job posting mentioned that you're looking for an analyst with experience in SQL, proficiency in a statistical programming language, and strong time management skills. During my previous role as a Data Entry Analyst, I had to multitask to balance the needs of the company (revenue) and the customer (service and quality standard). I've also been pursuing my passion for data, both at work and in my spare time. Over the past year, I've been able to achieve the following:

- Completed a case study using data set from Udemy to help increase revenue and identify the top performing courses
- Completed the Data Entry certification course on Entry Level with introduction to Tableau and Excel

I'm thrilled at the opportunity to use these experiences to fuel data-driven decisions at Chevron, and I'm keen to continue developing my skill set on the job. I am available for a Zoom call or in-person meeting to discuss how I can help Chevron with improving market product performance through data.

Thank you for your consideration,

James Brown, Data Analyst

Table of Contents

DATA REPORT ON COURSE REVENUE (UDEMY)

The problem

This is a report based on data collected from different subjects (Web Development, Business Finance, Musical Instruments & Graphics Design) and courses in each subject to understand where opportunities may lie to increase revenue and also track the performances of each course.

The Design

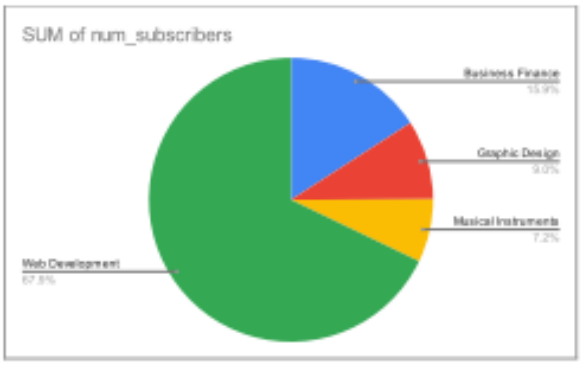
On receiving the data set and inputted into the spreadsheet, the data was cleaned using the Remove Duplicates functions, Find and Replace functions so as to have a consistent data format, Filter was introduced to remove empty cells. More information was needed so Date was extracted from the data set using Left/Right functions. IF function was used to determine the Free and Paid courses among each subject.

After getting all information ready and cleaned on the spreadsheet, sorting of the data set was done to determine the top 20 courses with the highest numbers of subscribers. The Advance sort range was used for this function and a new spreadsheet was created to know if there are courses that are beginner level and free. With the help of “IF AND” function this was determined and we only have two course that are free and also on beginner level. “VLOOKUP” function was used to move the date, numbers of subscribers, level, content duration from the main data set to this new spreadsheet where we have the top 20 most subscribed course. A Pivot table was then introduced to the main data set, we have our clean data to determine the summary of the information on the data set.

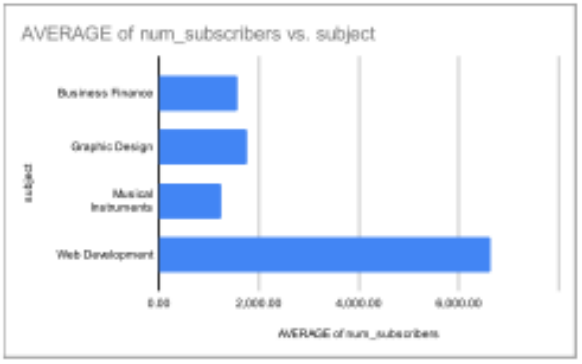
After creating a Pivot table from the clean data set, the data set was downloaded and the sheet where we have the 20 most subscribed courses. This was then introduced to tableau to get our visuals according to the pivot table created.

Kindly find below the visuals from the pivot table.

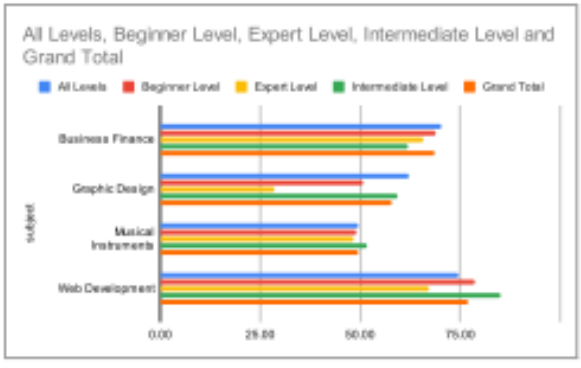
subject	SUM of num_sub
Business Financ	1868711
Graphic Design	1063148
Musical Instrume	845089
Web Developmen	7981935
Grand Total	**11760483**

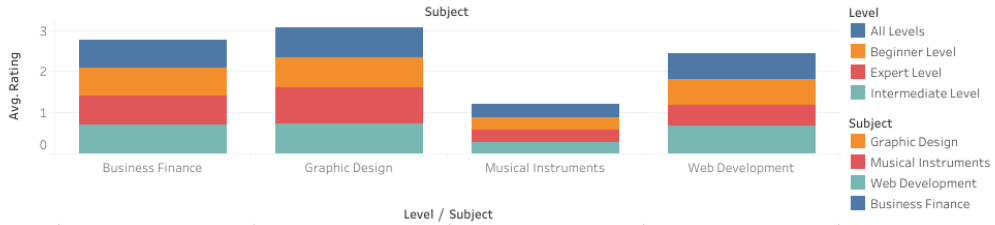


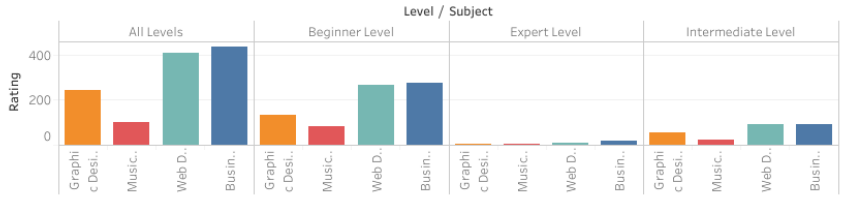
subject	AVERAGE of nu
Business Financ	1,569.03
Graphic Design	1,766.03
Musical Instrume	1,245.13
Web Developmen	6,635.02
Grand Total	**3,199.26**

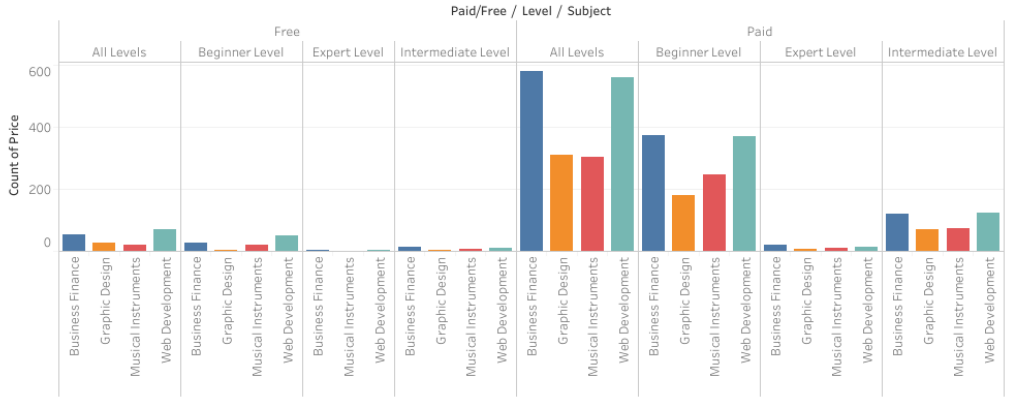


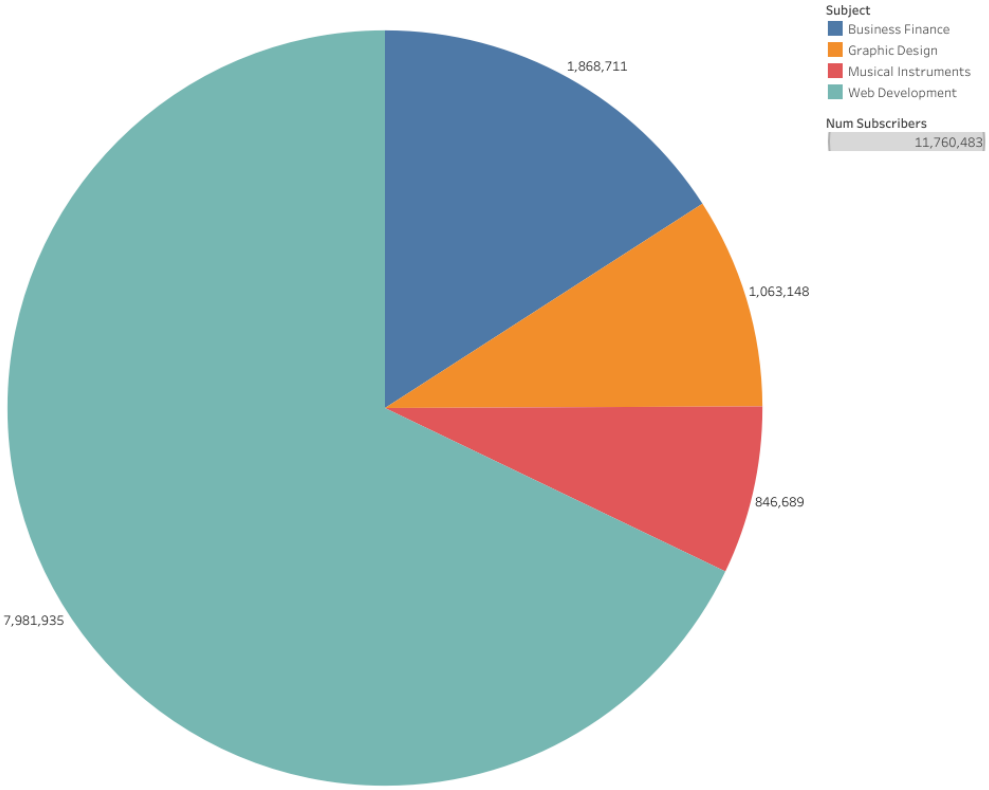
AVERAGE of pri level					
subject	All Levels	Beginner Level	Expert Level	Intermediate Lev	Grand Total
Business Financ	70.20	68.73	65.80	62.01	68.69
Graphic Design	62.12	50.68	28.57	59.41	57.89
Musical Instrume	49.58	48.98	48.33	51.60	49.56
Web Developmen	74.55	78.54	67.14	85.07	77.04
Grand Total	**66.75**	**65.24**	**58.62**	**66.94**	**66.12**




Subject
Avg. Rating
3
2
1
0
Business Finance
Graphic Design
Musical Instruments
Web Development
Level
All Levels
Beginner Level
Expert Level
Intermediate Level
Subject
Graphic Design
Musical Instruments
Web Development
Business Finance


Level / Subject
Rating
All Levels
Beginner Level
Expert Level
Intermediate Level
400
200
0
Graphi c Desi..
Music..
Web D..
Busin..


Paid/Free / Level / Subject
Count of Price
Free
Paid
All Levels
Beginner Level
Expert Level
Intermediate Level
600
400
200
0
Business Finance
Graphic Design
Musical Instruments
Web Development


Subject
Business Finance
Graphic Design
Musical Instruments
Web Development
Num Subscribers
11,760,483
1,868,711
1,063,148
846,689
7,981,935

ANALYSIS OF THE DATA

The root cause analysis was done by first framing the problem which is "To increase Revenue
and track the performance of each course" and after which the 5 WHY approach was used

1. Why do we want to increase revenue?
ANS: To be able to serve our subscribers very well with easy and smooth access to materials,
Qualified and respectable teachers and also to increase the profit of the organization.

2. Why should we increase the price of Web Development courses?
ANS: Because she feels web development courses are performing well above other subjects.

3. Why do we want to track the performance of each course?
ANS: To get more insight on how each course is performing and how much they generate for
the organization as a whole.

4. Why do we want to get more insight on each course performance?
ANS: So as to know the top performing courses and those to increase their prices as to generate
more revenue for the organization.

CONCLUSION

According to the insight and findings from the data set provided, I will be going with the Managers suggestion that Web development courses prices should be reviewed because amongst the 20 most subscribed courses, we have 17 courses on Web development. I will also suggest the free courses on web development courses on the top most subscribed courses should be paid courses. Lastly, I will suggest more courses on Web development to be introduced to the scheme. According to the findings also we have Business Finance to has more ratings than Web development and also more paid courses. I will suggest more advert be made for courses on Business Development which I think will help increase revenue.

REPORT ON SUPERMATKET BRANCHES SALES

This is a report based on data downloaded on “Kaggle” on Supermarket store branches sales analysis.

This data will be used to determine the differences between stores and why a store is performing better than the other, even though they are in the same area. This will also be used to determine how to increase profit for the store owners and improve sales among all stores.

The data set was cleaned and all duplicated removed with the help of remove duplicate function. The Filter was introduced to remove all empty cells. After the data has been cleaned and ready for analysis, I duplicated the data set into three segments to determine the top 20 stores with the highest number of daily customers, top 20 stores with the highest sales and lastly, top 20 stores with the highest number of items available. This is done with the help of Advance sort range.

VLOOKUP function was introduced to move information from the data set to a new spreadsheet where we have only the top 20 stores.

The if function was introduced to give the top 20 stores imaginary names instead of using the Store ID given in the data set.

A pivot table is then introduced for better comparison between stores. Kindly find below the visuals from the pivot table.

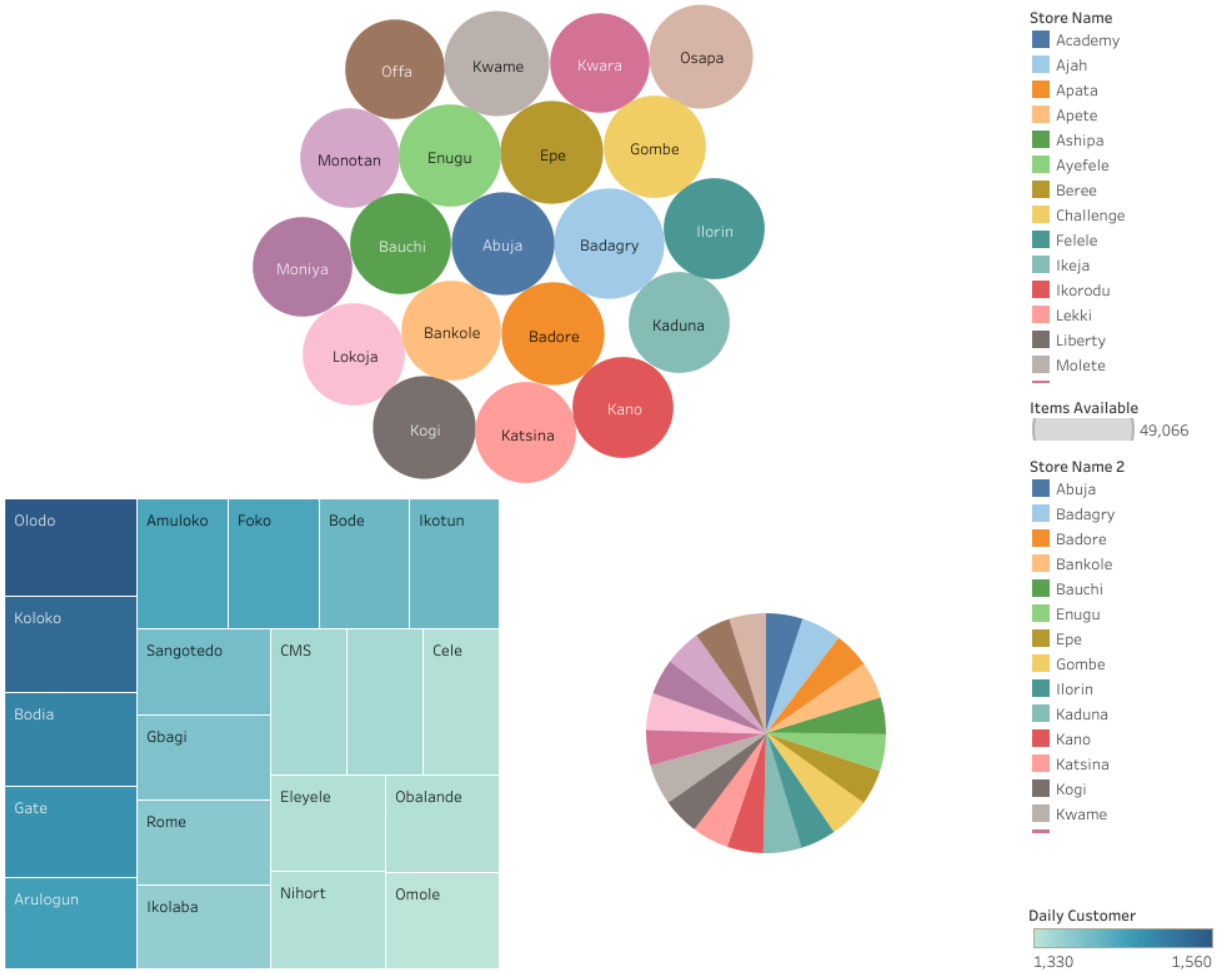
Store_Name	SUM of Items_A
Academy	2465
Ajah	2617
Apata	2414
Apete	2408
Ashipa	2474
Ayefele	2396
Beree	2384
Challenge	2667
Felele	2400
Ikeja	2492
Ikorodu	2397
Lekki	2493
Liberty	2420
Molete	2647
Oje	2385
Ojoo	2390
Omole	2396
Sango	2394
Taska	2436
UI	2391
Grand Total	**49066**

Store_Name	SUM of Daily_Cu
Amuloko	1440
Arulogun	1450
Bode	1410
Bodia	1490
Cele	1340
CMS	1350
Eleyele	1340
Foko	1440
Gate	1470
Gbagi	1390
Ikolaba	1370
Ikotun	1410
Koloko	1530
Nihort	1340
Obalande	1340
Olodo	1560
Omole	1330
Rome	1380
Sangotedo	1400
Stadium	1350
Grand Total	**28130**

Store_Name	SUM of Store_Sa
Abuja	101780
Badagry	116320
Badore	102310
Bankole	93950
Bauchi	97360
Enugu	98260
Epe	101820
Gombe	99480
Ilorin	95900
Kaduna	96610
Kano	96650
Katsina	97260
Kogi	100900
Kwame	105150
Kwara	94690
Lokoja	99570
Moniya	94370
Monotan	94170
Offa	94460
Osapa	102920
Grand Total	**1983930**

On the pivot table above, the first table represents the top 20 stores with the highest available items in each store where “Academy” has 2,465 total number of available items. The second table helps with the stores with the highest daily customers where we have “Amuloko” with 1,440 numbers of daily customers. While the last table presented us with the store with the highest sales where we have “Abuja” selling $101,780 in a month according to the data set given.

Kindly find the visuals from Tableau below. The monthly sales are visualized with packed burbles, the daily customers are visualized on the tree map while the available items are visualized on the pie chart.



Analysis

According to the visuals above we'll see that not one store is on all three set that is not one store has a high sale nor more customers walking in or high available items. This leaves us to no choice but to ask the following questions using the 5 WHY approach

- Why is the store with the highest number of available items not the store with the highest sales?
- Why do we have stores with high numbers of daily customers and not part of stores with highest sales?
- Why do we have stores with highest sales with less numbers of customers
- Why do we only have one store with the high available items and highest sales in a month?

Conclusion

After so much analysis of the data and comparison of the different stores, I will advise more data to be given on those stores with highest customers but low sales to take note of the customers to see if they are only window shopping and not buying also keep tab of products customers asks for but not available, probably they are not selling due to unavailability of products needed by the people in that vicinity.

I will also suggest more insight be given on the stores with the highest sales to keep tab of the products they sell more so as to introduce those products to other branches and also to be able to reduce the number of products in the stores with high available products and low sales.

We should also put into consideration the area of those stores with high sales and keep a tab on the type of customers that visits the store so as to know the type of advertisement to be done for other stores with low sales also to note the prices of those products that sells the most maybe pricing is the reason why the other stores have more products in their inventory with low sales.

Appendix

Data set used for the Udemy Project
udemy data set

Data set used for the Capstone Project
Capstone project