

## KEY INSIGHT FROM CLEANED DATASET

HI, MY NAME IS IFEOLUWA SAMUEL SUNDAY, AN ASPIRING DATA SCIENTST STARTING WITH DATA ANALYTICS, FIRST OF ALL. MY TASK IS TO CLEAN MY CLIENT DATASET NAMED 'SHOPSMART' AS THE NAME IMPLIES. IT DATA NEED TO BE CORRECT AND SMART TOO. LO

THIS OVERALL DATASET REQUIRED ME TO FIRST STUDY WHAT THE DATA IS ALL ABOUT UNDERTSANDING THE JOINED DATASETS TABLE AND BEING ABLE TO VISUALLY PRESENT THIS INFORMATION.

~ I UNDERSTOOD

~ CLEANED

~ GROUPED

~ VISUALISE USING PYTHON(ONE THING I FIND IS ALL THIS CAN BE DONE USING EXCEL TOO, BUT I THINK IT GETS CRAZIER WHEN IT GETS TO LARGE DATASETS EXCEL CANNOT HANDLE.)

SO FIRST, MY TASK IS TO SHOW SALES PERFORMANCE, CUSTOMER PATTERNS AND PRODUCT TRENDS.

THIS DATASET CONSIST OF CUSTOMERS INFO AND THEIR PURCHASES.

STARTED BY VIEWING THE INFO FOR BETER UNDERTSANDING

- (*HEAD,INFO,DESCRIBE*)
- DROPPED THE DUPLICATES
- MERGED BOTHY DATASETS
- CREATE NEW COLUMNS{TOTAL AND CHANGED THE DATE TO STANDARD ONE 01-01-2001 –JANUARY },
- PROCEEDED TO FILTERING OUT THE TOP CUSTOMERS BY USING GROUPBY
- SORTED IT BY ACSENDING
- THEN GROUPED BY CATEGORY LIKE PRODUCT CATEGORY AND IT TOTAL PRICE
- ALSO DID TOTAL PRICE BY CITY
- THEN LASTLY TOP PRODUCT

AFTER SUCCESFULLY FILTERING OUT THE RELEVANT INFORMATION FROM OUR DATA, I PROCEEDED TO VISUALISING THEM TO MAKE IT EASY FOR NON NUMERICAL STAKEHOLDERS TO BE ABLE TO UNDERSTAND BETTER.

### BELOW ARE MY KEY INSIGHTS

- NOW KEY INSIGHT ON **MONTHLY SALES**  
HIGH SALES IN JANUARY BUT HUGE SALES DECLINE IN THE FOLLOWING MONTH  
MAYBE A HIGH RISE IN PRICE.
- ELECTRONICS DOMINATED AND OBTAINED A VERY HIGH SALES COMPARE TO FURNITURE  
LOCATION MIGHT MATTER IN THIS SITUATION
- PRICE DISTRIBUTION ROSE UP AROUND 200 X 3 ORDERS, MAYBE MOST CUSTOMERS  
WHO MADE THE MOST PURCHASE BOUGHT 2 X 200 EACH
- THE RELATIONSHIP BETWEEN QUANTITY AND SALES KINDA HELP BUSINESS TO MAKE  
INFORMED DECISION ON DEMAND AND SUPPLY
- AND LASTLY TOP CUSTOMER PURCHASES, THE TOP 3 OUT OF FIVE CUSTOMER WHO  
PURCHASED THE MOST PRODUCT ACTUALLY PURCHASED ELECTRONICS WHICH IS HIGHER  
THAN THE AVERAGE OF THE 5 CUSTOMERS. MEANING ELECTRONICS ALWAYS  
DOMINATE. MORE FOCUS ON FURNITURE TOO MAYBE ADVERTISE MORE OF FURNITURE  
OR GIVE DISCOUNTS TO MAKE MORE SALES.

OVERALL THIS PROJECT KINDA OPEN MORE OF MY UNDERSTANDING OF HOW DATA ANALYTICS WORKS. THE CLEANING PROCESS AND EXPECIALLY THE VISUALIZATION. I STILL DO NOT HAVE MOST OF THE CODE IN MY HEAD THOUGH. BUT I CAN NOW UNDERSTAND OR HAVE AN IDEA OF HOW THEY WORK NOW.

### REFLECTION QUESTION ANSWERS

#### **Reflection Questions ( Answers)**

1. What new trend did you discover in this project? THE OVERALL TRENDS I VISUALISE
2. How did filtering and grouping help make sense of the data? IT SORT OUT RELEVANT DATA NEEDED

3. How did visualization change how you understood the results? BROUGHT MORE CLARITY TO THE DATA IN REAL TIME
4. What would you do next if you were the ShopSmart analyst? COMMUNICATE INSIGHT THAT COULD HELP THE STORE HAVE MORE SALES AND EXCAPE DIPS