CS544 Final Project

Amazon Sale Report Analysis

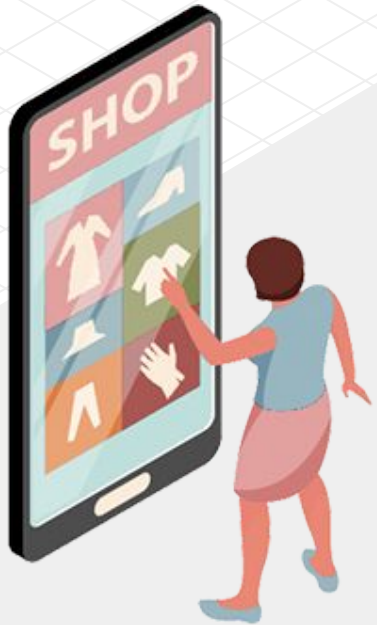
Shou-Tzu Han

TABLE OF CONTENTS

01

Preparing the data

1) Import the data set into R.
2) Document the steps for the import process and data preparation, processing, and cleaning procedures had to be done. Any R code used in the process should be included.



Preparing the data

1. Install the tidyverse package.
2. Load the tidyverse library.
3. Import the data from the "Amazon Sale Report.csv" file.
4. Remove rows with missing values.
5. Filter rows where the 'Amount' is greater than 100.

https://www.kaggle.com/datasets/thedevastator/unlock-profits-with-e-commerce-sales-data?resource=download

Before cleaning the data

index	Order ID	Date	Status	Fulfilment	Sales Channel	ship-service-level	Style	SKU	Category	...	currency	Amount	ship-city	ship-state	ship-postal-code	ship-country	promotion-ids	B2B	fulfilled-by	Unnamed: 22
<dbl>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	...	<chr>	<dbl>	<chr>	<chr>	<dbl>	<chr>	<chr>	<lgl>	<chr>	<lgl>
0	405-8078784-5731545	04-30-22	Cancelled	Merchant	Amazon.in	Standard	SET389	SET389-KR-NP-S	Set	...	INR	647.62	MUMBAI	MAHARASHTRA	400081	IN		NA	FALSE	Easy Ship	NA
1	171-9198151-1101146	04-30-22	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	JNE3781	JNE3781-KR-XXXL	kurta	...	INR	406.00	BENGALURU	KARNATAKA	560085	IN	Amazon PLCC Free-Financing Universal Merchant AAT-WNKTBO3K27EJC,Amazon PLCC Free-Financing Universal Merchant AAT-QX3UCCJESKPA2,Amazon PLCC Free-Financing Universal Merchant AAT-5QQ7BIYYQEDN2,Amazon PLCC Free-Financing Universal Merchant AAT-DSJ2QRXXWXVMQ,Amazon PLCC Free-Financing Universal Merchant AAT-CXJHMC2YJUK76,Amazon PLCC Free-Financing Universal Merchant AAT-CC4FAVTYR4X7C,Amazon PLCC Free-Financing Universal Merchant AAT-FYWJCQHU2XQ2U,Amazon PLCC Free-Financing Universal Merchant AAT-EOKPWFWYW7Y6I,Amazon PLCC Free-Financing Universal Merchant AAT-CXNSLNBROFDW6,Amazon PLCC Free-Financing Universal Merchant AAT-R7GXNZWISTRFA,Amazon PLCC Free-Financing Universal Merchant AAT-WSJLDN3X7KEMO,Amazon PLCC Free-Financing Universal Merchant AAT-VL6FGQVGQVXUS,Amazon PLCC Free-Financing Universal Merchant AAT-EOKPWFWYW7Y6I,Amazon PLCC Free-Financing Universal Merchant AAT-ZYL5UPUNW6T62,Amazon PLCC Free-Financing Universal Merchant AAT-XVPICCHRWDCAI,Amazon PLCC Free-Financing Universal Merchant AAT-ETXQ3XXWMRXBG,Amazon PLCC Free-Financing Universal Merchant AAT-7X3XCTYGB4VBE,Amazon PLCC Free-Financing Universal Merchant AAT-HGD3WTS3MHM,Amazon PLCC Free-Financing Universal Merchant AAT-26ZDKNME27X42,Amazon PLCC Free-Financing Universal Merchant AAT-4ZF5KN6E4LJK4,Amazon PLCC Free-Financing Universal Merchant AAT-7RCXIKUAX7DPY,Amazon PLCC Free-Financing Universal Merchant AAT-BRSZZ45H6MHAQ,Amazon PLCC Free-Financing Universal Merchant AAT-MKLXOOZWQL7GO,Amazon PLCC Free-Financing Universal Merchant AAT-CB7UNXEXGIJTC,Amazon PLCC Free-Financing Universal Merchant #MP-gzasho-1593152694811,Amazon PLCC Free-Financing Universal Merchant AAT-WLBA4GZ52EAH4	FALSE	Easy Ship	NA
2	404-0687676-7273146	04-30-22	Shipped	Amazon	Amazon.in	Expedited	JNE3371	JNE3371-KR-XL	kurta	...	INR	329.00	NAVI MUMBAI	MAHARASHTRA	410210	IN	IN Core Free Shipping 2015/04/08 23-48-5-108	TRUE	NA	NA
3	403-9615377-8133951	04-30-22	Cancelled	Merchant	Amazon.in	Standard	J0341	J0341-DR-L	Western Dress	...	INR	753.33	PUDUCHERRY	PUDUCHERRY	605008	IN		NA	FALSE	Easy Ship	NA
4	407-1069790-7240320	04-30-22	Shipped	Amazon	Amazon.in	Expedited	JNE3671	JNE3671-TU-XXXL	Top	...	INR	574.00	CHENNAI	TAMIL NADU	600073	IN		NA	FALSE	NA	NA
5	404-1490984-4578765	04-30-22	Shipped	Amazon	Amazon.in	Expedited	SET264	SET264-KR-NP-XL	Set	...	INR	824.00	GHAZIABAD	UTTAR PRADESH	201102	IN	IN Core Free Shipping 2015/04/08 23-48-5-108	FALSE	NA	NA

After cleaning the data

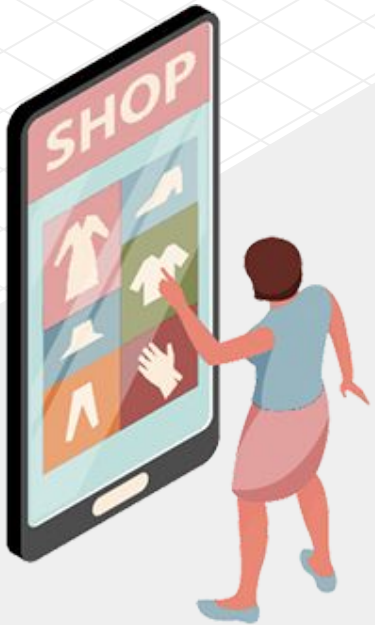
Missing value

index	Order ID	Date	Status	Fulfilment	Sales Channel	ship-service-level	Style	SKU	Category	...	Amount	ship-city	ship-state	ship-postal-code	ship-country	promotion-ids	B2B	fulfilled-by	Unnamed: 22	Amount_in_USD
<dbl>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	...	<dbl>	<chr>	<chr>	<dbl>	<chr>	<chr>	<lgl>	<chr>	<lgl>	<dbl>
49051	408-4858463-2356347	05-31-22	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	J0385	J0385-KR-M	kurta	...	888	RAJAHMUNDRY	ANDHRA PRADESH	533126	IN	Amazon PLCC Free-Financing Universal Merchant AAT-WNKTBO3K27EJC,Amazon PLCC Free-Financing Universal Merchant AAT-QX3UCCJESKPA2,Amazon PLCC Free-Financing Universal Merchant AAT-5QQ7BIYYQEDN2,Amazon PLCC Free-Financing Universal Merchant AAT-DSJ2QRXXWXVMQ,Amazon PLCC Free-Financing Universal Merchant AAT-CXJHMC2YJUK76,Amazon PLCC Free-Financing Universal Merchant AAT-CC4FAVTYR4X7C,Amazon PLCC Free-Financing Universal Merchant AAT-FYWJCQHU2XQ2U,Amazon PLCC Free-Financing Universal Merchant AAT-EOKPWFWYW7Y6I,Amazon PLCC Free-Financing Universal Merchant AAT-URPEBMPS7TSPE,VPC-44571-41445473 Coupon,Amazon PLCC Free-Financing Universal Merchant AAT-5ZEB2OGNZ32MC,Amazon PLCC Free-Financing Universal Merchant AAT-CAK4QZ7KEFRQU,Amazon PLCC Free-Financing Universal Merchant AAT-YZJSBRTPINRLK,Amazon PLCC Free-Financing Universal Merchant AAT-SNWJTEAZA2JRI,Amazon PLCC Free-Financing Universal Merchant AAT-BBJ6FVTP3QKBO,Amazon PLCC Free-Financing Universal Merchant AAT-TNYBHEAO7HOHO,Amazon PLCC Free-Financing Universal Merchant AAT-BRSZZ45H6MHAQ,Amazon PLCC Free-Financing Universal Merchant AAT-MKLXOOZWQL7GO,Amazon PLCC Free-Financing Universal Merchant AAT-CB7UNXEXGIJTC,Amazon PLCC Free-Financing Universal Merchant #MP-gzasho-1593152694811,Amazon PLCC Free-Financing Universal Merchant AAT-WLBA4GZ52EAH4,Amazon PLCC Free-Financing Universal Merchant AAT-QMFZ3M7T5NPDS,Amazon PLCC Free-Financing Universal Merchant #MP-pmilan-1601980554292	FALSE	Easy Ship	FALSE	976.8
49077	402-7697834-2682766	05-31-22	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	JNE3654	JNE3654-TP-XXL	Top	...	371	CHENNAI	TAMIL NADU	600042	IN	Amazon PLCC Free-Financing Universal Merchant AAT-WNKTBO3K27EJC,Amazon PLCC Free-Financing Universal Merchant AAT-QX3UCCJESKPA2,Amazon PLCC Free-Financing Universal Merchant AAT-5QQ7BIYYQEDN2,Amazon PLCC Free-Financing Universal Merchant AAT-DSJ2QRXXWXVMQ,Amazon PLCC Free-Financing Universal Merchant AAT-CXJHMC2YJUK76,Amazon PLCC Free-Financing Universal Merchant AAT-CC4FAVTYR4X7C,Amazon PLCC Free-Financing Universal Merchant AAT-FYWJCQHU2XQ2U,Amazon PLCC Free-Financing Universal Merchant AAT-EOKPWFWYW7Y6I,Amazon PLCC Free-Financing Universal Merchant AAT-CAK4QZ7KEFRQU,Amazon PLCC Free-Financing Universal Merchant AAT-YZJSBRTPINRLK,Amazon PLCC Free-Financing Universal Merchant AAT-SNWJTEAZA2JRI,Amazon PLCC Free-Financing Universal Merchant AAT-BBJ6FVTP3QKBO,Amazon PLCC Free-Financing Universal Merchant AAT-BRSZZ45H6MHAQ,Amazon PLCC Free-Financing Universal Merchant AAT-MKLXOOZWQL7GO,Amazon PLCC Free-Financing Universal Merchant AAT-CB7UNXEXGIJTC,Amazon PLCC Free-Financing Universal Merchant #MP-gzasho-1593152694811,Amazon PLCC Free-Financing Universal Merchant AAT-WLBA4GZ52EAH4,Amazon PLCC Free-Financing Universal Merchant #MP-pmilan-1601980554292	FALSE	Easy Ship	FALSE	408.1
49081	402-7031777-0396341	05-31-22	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	JNE3798	JNE3798-KR-XS	Western Dress	...	771	NEDUMANGAD	KERALA	695541	IN	Amazon PLCC Free-Financing Universal Merchant AAT-WNKTBO3K27EJC,Amazon PLCC Free-Financing Universal Merchant AAT-QX3UCCJESKPA2,Amazon PLCC Free-Financing Universal Merchant AAT-5QQ7BIYYQEDN2,Amazon PLCC Free-Financing Universal Merchant AAT-DSJ2QRXXWXVMQ,Amazon PLCC Free-Financing Universal Merchant AAT-CXJHMC2YJUK76,Amazon PLCC Free-Financing Universal Merchant AAT-CC4FAVTYR4X7C,Amazon PLCC Free-Financing Universal Merchant AAT-FYWJCQHU2XQ2U,Amazon PLCC Free-Financing Universal Merchant AAT-EOKPWFWYW7Y6I,Amazon PLCC Free-Financing Universal Merchant AAT-URPEBMPS7TSPE,Amazon PLCC Free-Financing Universal Merchant AAT-CAK4QZ7KEFRQU,Amazon PLCC Free-Financing Universal Merchant AAT-YZJSBRTPINRLK,Amazon PLCC Free-Financing Universal Merchant AAT-SNWJTEAZA2JRI,Amazon PLCC Free-Financing Universal Merchant AAT-BBJ6FVTP3QKBO,Amazon PLCC Free-Financing Universal Merchant AAT-BRSZZ45H6MHAQ,Amazon PLCC Free-Financing Universal Merchant AAT-MKLXOOZWQL7GO,Amazon PLCC Free-Financing Universal Merchant AAT-CB7UNXEXGIJTC,Amazon PLCC Free-Financing Universal Merchant #MP-gzasho-1593152694811,Amazon PLCC Free-Financing Universal Merchant AAT-WLBA4GZ52EAH4	FALSE	Easy Ship	FALSE	848.1
49083	403-3845187-4421165	05-31-22	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	J0157	J0157-DR-XL	Western Dress	...	690	PIMPRI CHINCHWAD	MAHARASHTRA	411027	IN	Amazon PLCC Free-Financing Universal Merchant AAT-WNKTBO3K27EJC,Amazon PLCC Free-Financing Universal Merchant AAT-QX3UCCJESKPA2,Amazon PLCC Free-Financing Universal Merchant AAT-5QQ7BIYYQEDN2,Amazon PLCC Free-Financing Universal Merchant AAT-DSJ2QRXXWXVMQ,Amazon PLCC Free-Financing Universal Merchant AAT-CXJHMC2YJUK76,Amazon PLCC Free-Financing Universal Merchant AAT-SL3Y2YRMAUNFC,Amazon PLCC Free-Financing Universal Merchant AAT-CC4FAVTYR4X7C,Amazon PLCC Free-Financing Universal Merchant AAT-FYWJCQHU2XQ2U,Amazon PLCC Free-Financing Universal Merchant AAT-WSJLDN3X7KEMO,Amazon PLCC Free-Financing Universal Merchant AAT-VL6FGQVGQVXUS,Amazon PLCC Free-Financing Universal Merchant AAT-EOKPWFWYW7Y6I,Amazon PLCC Free-Financing Universal Merchant AAT-URPEBMPS7TSPE,Amazon PLCC Free-Financing Universal Merchant AAT-5VXFLMD6EKIFC,Amazon PLCC Free-Financing Universal Merchant AAT-CAK4QZ7KEFRQU,Amazon PLCC Free-Financing Universal Merchant AAT-YZJSBRTPINRLK,Amazon PLCC Free-Financing Universal Merchant AAT-SNWJTEAZA2JRI,Amazon PLCC Free-Financing Universal Merchant AAT-BBJ6FVTP3QKBO,Amazon PLCC Free-Financing Universal Merchant AAT-BRSZZ45H6MHAQ,Amazon PLCC Free-Financing Universal Merchant AAT-MKLXOOZWQL7GO,Amazon PLCC Free-Financing Universal Merchant AAT-CB7UNXEXGIJTC,Amazon PLCC Free-Financing Universal Merchant #MP-gzasho-1593152694811,Amazon PLCC Free-Financing Universal Merchant AAT-WLBA4GZ52EAH4	FALSE	Easy Ship	FALSE	759.0

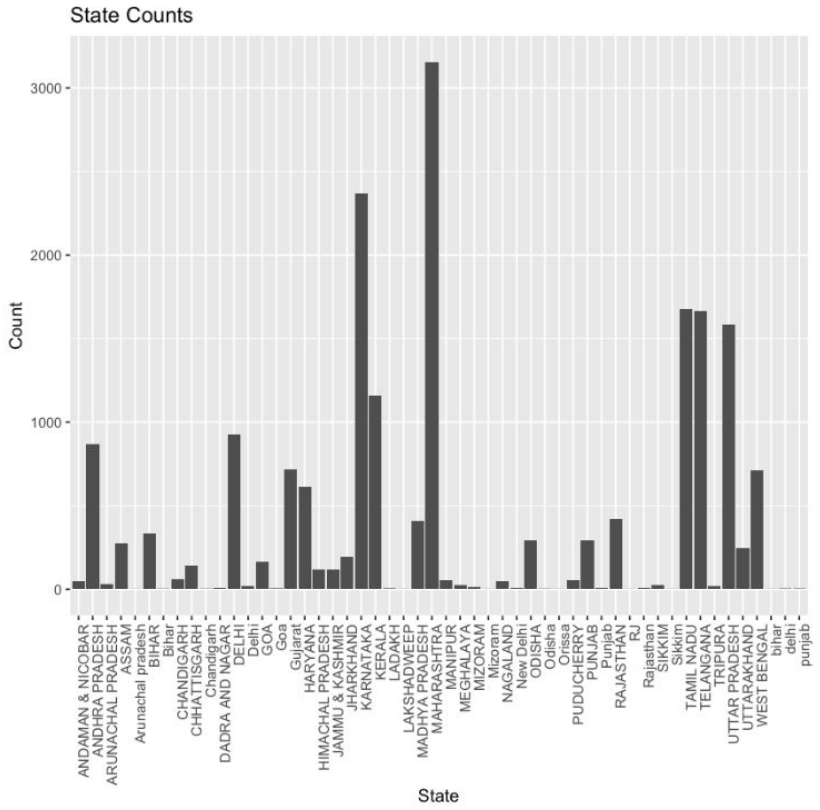
02

Analyzing the data

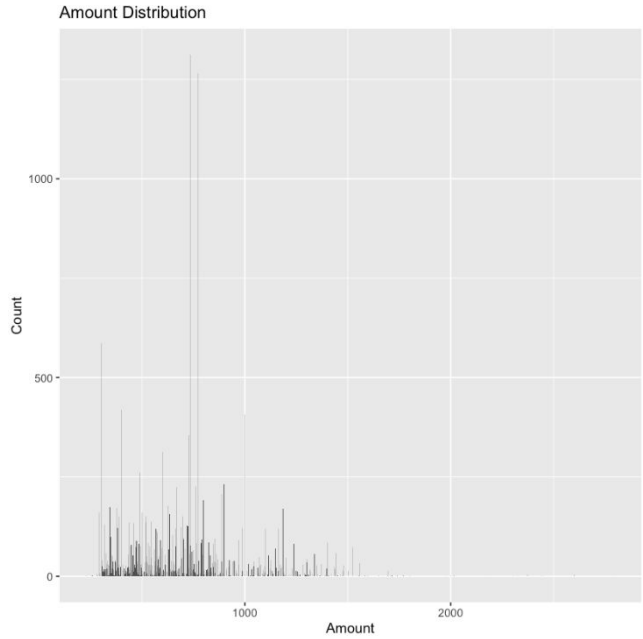
1) Analyzing categorical and numerical variables
2) Analyzing sets of two or more variables
3) Demonstrating the Central Limit Theorem
4) Applying various sampling methods
5) Additional feature implementation



“State” Variable Analysis



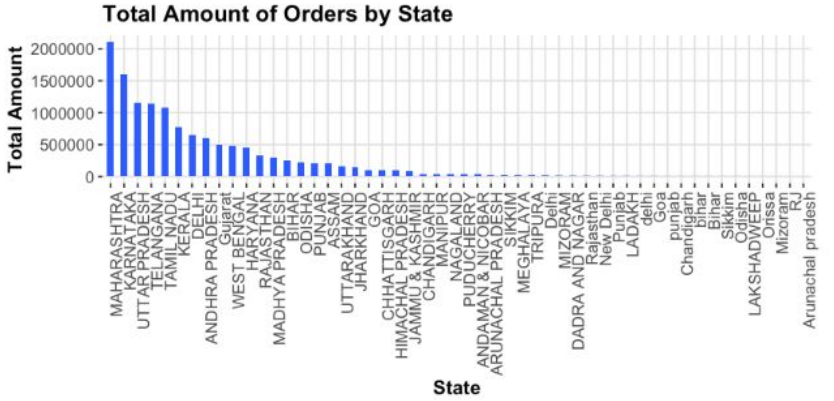
“Amount” Variable Analysis

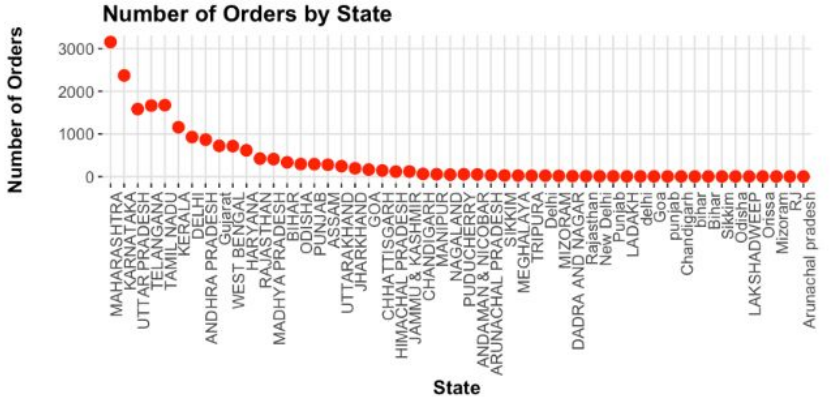


Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
229.0	486.0	692.0	689.5	788.0	2796.0

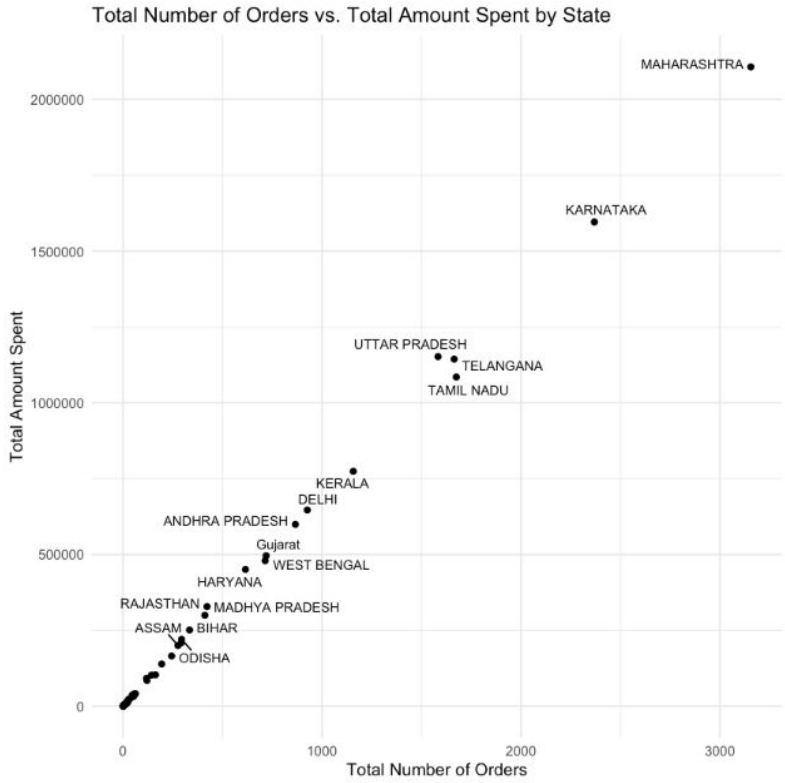
Analysis of Two Variables

ship-state	total_amount	total_orders
<chr>	<dbl>	<int>
MAHARASHTRA	2106917	3155
KARNATAKA	1595847	2369
UTTAR PRADESH	1152335	1583
TELANGANA	1143765	1664
TAMIL NADU	1084697	1675





Relationship between “# of order” and “amount spend by state”



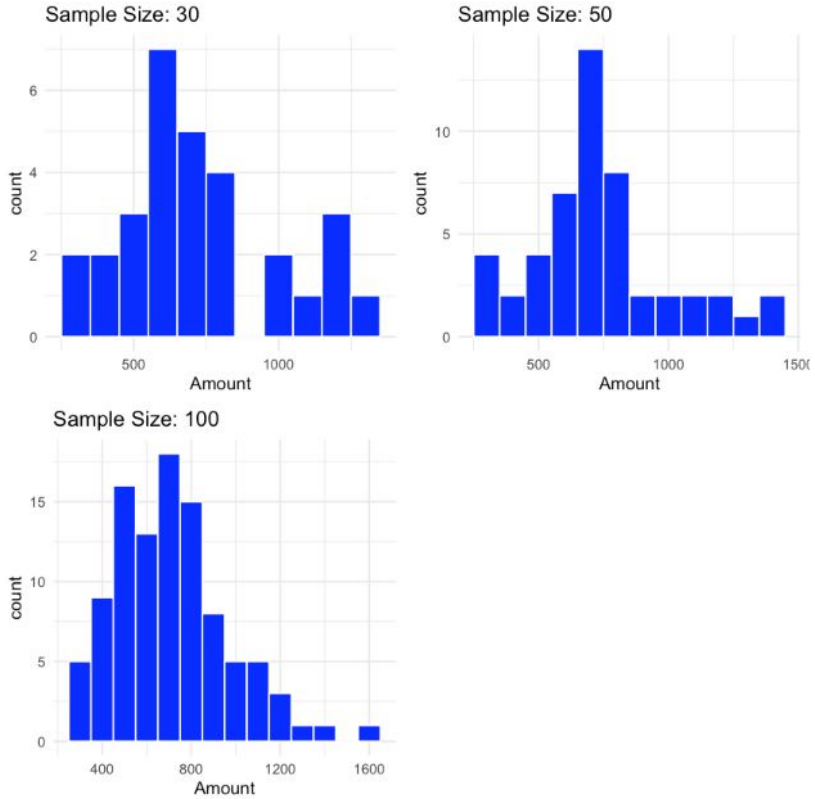
Conclusion:

Certain states exhibit higher levels of spending on Amazon orders compared to others

Marketing strategies:

Allocate resources more effectively to target regions with higher consumer spending.

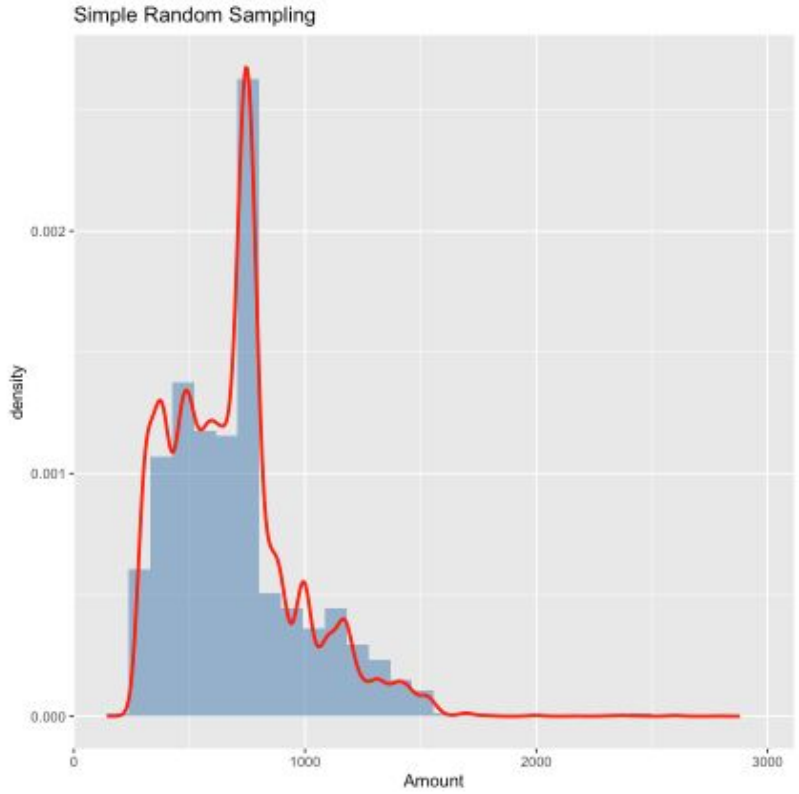
Random samples of “Amount”

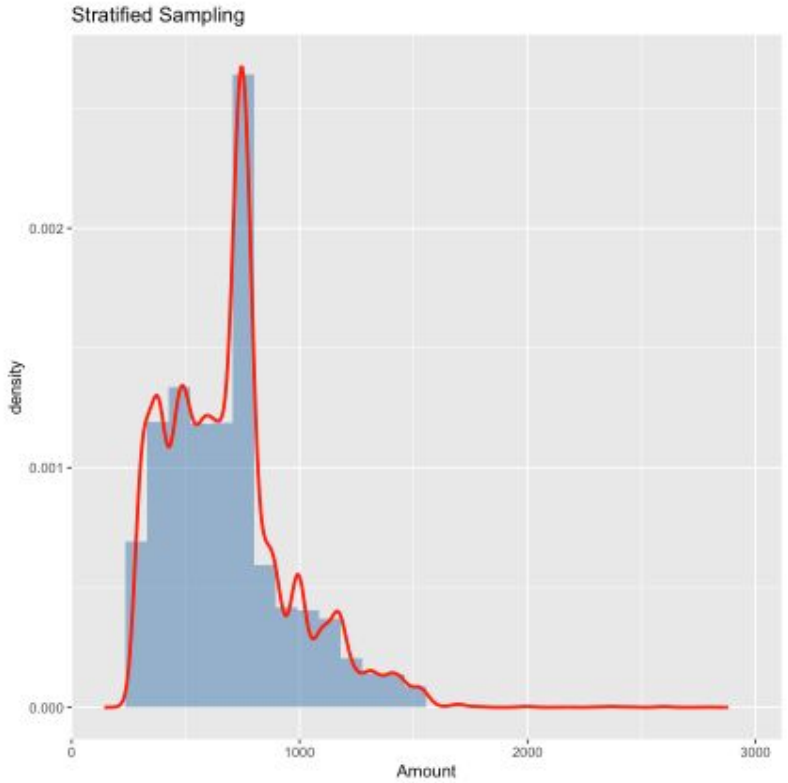


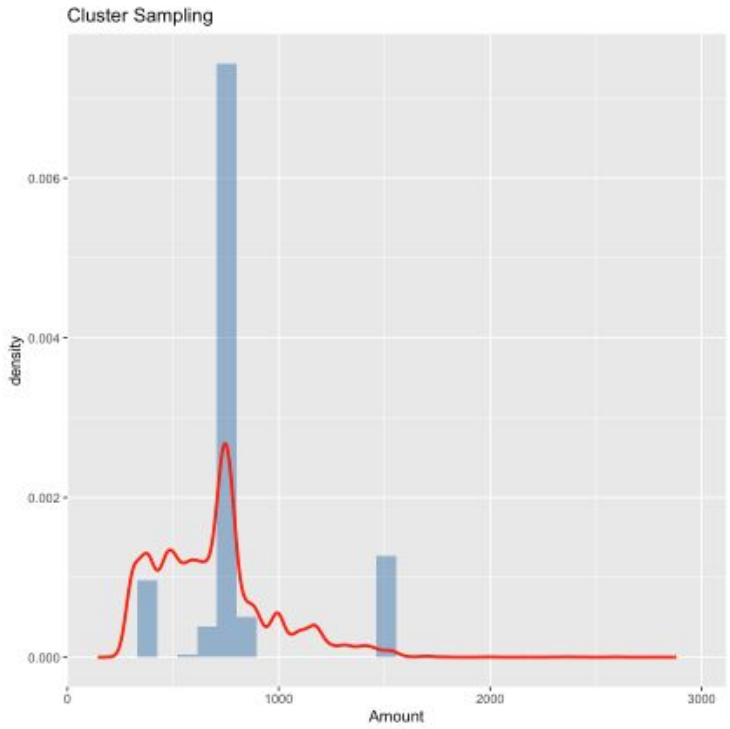
Conclusion:
The larger the number of samples selected, the higher the accuracy of the inference of the statistical analysis results to the actual situation.

```
     Amount
Min.   : 229.0
1st Qu.: 486.0
Median : 692.0
Mean   : 689.5
3rd Qu.: 788.0
Max.   :2796.0
```

3 sampling methods on “Amount”





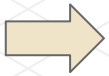
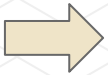


Stratified sampling provide more accurate estimates for each subgroup, and is more similar to our amount variable distribution

```
     Amount
 Min.   : 229.0
 1st Qu.: 486.0
 Median : 692.0
 Mean   : 689.5
 3rd Qu.: 788.0
 Max.   :2796.0
```

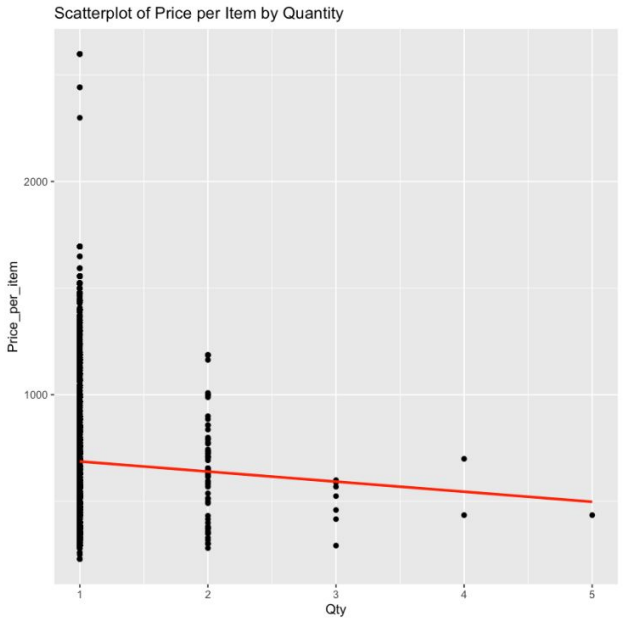
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Simple Random Sampling	229.0	487.0	699.0	705.9	799.0	2442.0
Stratified Sampling	280.0	486.0	690.0	683.5	786.0	1695.0
Cluster Sampling	348.0	799.0	799.0	847.2	799.0	1556.0

Conclusion

Simple Random Sampling		when you want to obtain a representative sample of the entire dataset
Stratified Sampling		when you want to ensure representation from different subgroups within the dataset
Cluster Sampling		when it is costly or impractical to sample individual units within each cluster

Additional Analysis: correlation between Qty and Price_per_item

Hyphosis: whether buy more item, per item price decrease.



Conclusion:

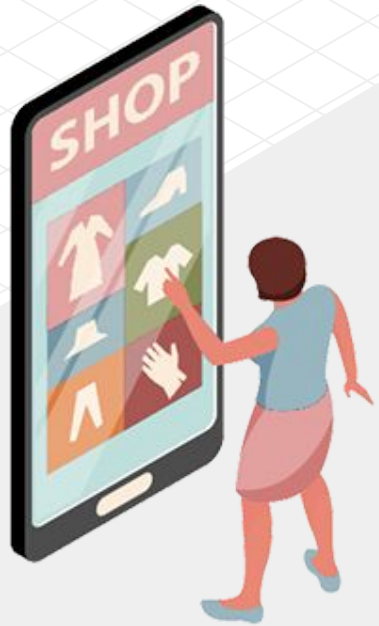
Customers purchase larger quantities, they can enjoy a lower price per item

Marketing strategies:

This could be attributed to volume discounts or promotional offers that are commonly applied when customers buy in bulk.

03

Q&A



THANKS

If you like this project, please give it a STAR on my GitHub. Thank you!

https://github.com/52147/Amazon-Sales-Report-Analysis