

Mariano Méndez-Suárez

DESCRIPTIVE  
AND INFERENTIAL  
STATISTICS

PROBLEMS  
AND SOLUTIONS



# **Descriptive and Inferential Statistics**

Problems and Solutions

Madrid, 2025

Mariano Méndez-Suárez

# **Descriptive and Inferential Statistics**

Problems and Solutions

April, 2025

*Descriptive and Inferential Statistics: Problems and Solutions*

Mariano Méndez-Suárez

All rights reserved.

Any form of reproduction, distribution, public communication or transformation of this work may only be carried out with the authorization of its owners, with the exceptions provided for by law.

Please contact CEDRO (Spanish Centre for Reprographic Rights) if you need to photocopy or scan any part of this work ([www.cedro.org](http://www.cedro.org)).

© 2025, ESIC Editorial  
Avda. de Valdenigrales, s/n  
28223 Pozuelo de Alarcón (Madrid)  
Tel.: 91 452 41 00  
[www.esic.edu/editorial](http://www.esic.edu/editorial)  
@EsicEditorial

ISBN: 978-84-1192-155-8  
Legal Deposit: M-5076-2025

Cover: Zita Moreno Puig  
Design: Santiago Díez Escribano  
Printed by: Gráficas Dehon

A publication by



Printed in Spain

*This book has been printed with ecological ink and sustainable paper.*

To Pilar

# Contents

Chapter 1. <b>Introduction to Descriptive Statistics</b> .....	19
1.1. Key Concepts and Techniques for Frequency Tables .....	19
Frequency Table Overview .....	19
Key Calculations .....	20
Example Step-by-Step Frequency Table Calculation .....	20
Step-by-Step Solution .....	22
1.2. Exercises on Descriptive Statistics .....	24
Exercise 1: Daily Visits to a Website.....	24
Exercise 2: Number of Products Sold per Day (Marketing) .....	26
Exercise 3: Customer Reviews per Day (Marketing) .....	27
Exercise 4: Daily Transactions in a Business .....	28
Exercise 5: Units Produced per Day (Business) .....	30
1.3. Proposed Exercises .....	31
Exercise 1: Daily Visits to an E-commerce Website.....	31
Exercise 2: Number of Products Sold per Week (Marketing) .....	32
Exercise 3: Customer Support Tickets per Day (Digital Business) ....	32
Exercise 4: Monthly Subscriptions to a Digital Service .....	33
Exercise 5: Daily Transactions in an Online Store.....	33
Chapter 2. <b>Measures of Central Tendency</b> .....	35
2.1. Central tendency measures .....	35
2.2. Exercises on Mean, Median, Mode, and Percentiles .....	36
Exercise 1: Customer Daily Screen Time on a news webpage .....	36
Exercise 2: Customer Daily Screen Time on a sports App .....	37

Exercise 3: Customer Daily Screen Time on an App .....	38
Exercise 4: Number of Orders Processed per Day in an Online Store	39
Exercise 5: Daily User Sign-ups on a Digital Platform.....	41
2.3. Proposed Exercises on Mean, Median, Mode, and Percentiles.....	42
Exercise 1: Daily Visits to a Marketing Blog.....	42
Exercise 2: Weekly Sales of Products in an Online Store .....	43
Exercise 3: Monthly Subscriptions to a Digital Service .....	43
Exercise 4: Daily Revenue of an E-commerce Site (Digital Business) ..	44
Exercise 5: Weekly Active Users on a Social Media Platform (Digital Business) .....	44
2.4. Measures of Shape: Skewness and Kurtosis.....	45
Skewness.....	45
Kurtosis.....	46
Examples .....	47
Chapter 3. <b>Concentration Measures: Lorenz Curve and Gini Index</b> .....	49
3.1. Lorenz Curve .....	49
Key Points .....	50
3.2. Gini Index.....	50
Applications .....	50
Example 1: Income Distribution among a Population .....	50
Example 2: Daily Sales in a Retail Store.....	53
Example 3: Monthly Marketing Campaign Expenses.....	55
Example 4: Weekly Product Sales in an E-commerce Store .....	57
Example 5: Monthly Revenue from Digital Services .....	60
Example 6: Quarterly Profits of a Marketing Firm .....	62
Example 7: Yearly Sales Revenue of a Digital Product.....	64
Chapter 4. <b>Measures of Dispersion: Variance and Standard Deviation</b> ...	67
4.1. Calculating Variance and Standard Deviation .....	67
4.2. Variance Calculation when each Data Point has a Frequency of 1....	68
Exercise 1: Daily Revenue from a Marketing Campaign.....	68
Exercise 2: Weekly Product Sales in a Business.....	70
Exercise 3: Monthly Subscriptions to a Digital Service .....	71
Exercise 4: Daily Active Users on a Social Media Platform .....	72
Exercise 5: Weekly Revenue from an E-commerce Website .....	73
4.3. Exercises on Measures of Dispersion when each Data Point has a Frequency greater than 1 .....	75

Exercise 1: Analysis of Sales for a Digital Product.....	75
Exercise 2: Click-Through Rate (CTR) in Advertising Campaigns ....	76
Exercise 3: Social Media Engagement Rates .....	77
Exercise 4: Monthly Revenue from Subscriptions .....	78
Exercise 5: Customer Retention Rates .....	79
4.4. Proposed Exercises on Measures of Dispersion with absolute frequency equal to 1 .....	80
Exercise 1: Monthly Sales of a Product .....	80
Exercise 2: Daily Visitors to a Website .....	80
Exercise 3: Weekly Earnings from a Subscription Service .....	81
Exercise 4: Monthly Active Users on an App .....	81
Exercise 5: Quarterly Revenue of a Business .....	82
4.5. Proposed Exercises with absolute frequency greater than 1 .....	82
Exercise 1: Conversion Rate of a Campaign .....	82
Exercise 2: Customer Ratings for a New Product .....	83
Exercise 3: Ad Impressions .....	83
Exercise 4: Page Views .....	83
Exercise 5: Customer Purchase Frequency .....	84
Chapter 5. <b>Index Numbers</b> .....	85
5.1. Index numbers at a glance .....	85
5.2. Simple Indexes .....	85
5.3. Aggregate and Weighted Indexes .....	86
Aggregate Indexes .....	86
Weighted Indexes .....	86
5.4. Laspeyres Index $L_0^t$ .....	86
Definition.....	86
Practical use.....	87
Example Calculation.....	87
5.5. Paasche Index $P_0^t$ .....	87
Definition.....	87
Practical use.....	88
Example Calculation.....	88
5.6. Comparison of Laspeyres and Paasche Indexes .....	88
5.7. How the Consumer Price Index (CPI) is Calculated in Spain .....	89
Steps for Calculating CPI in Spain .....	89
Example of CPI Calculation .....	90
5.8. How the NASDAQ Index is Calculated .....	90

Type of Index Used.....	90
Steps for Calculating the NASDAQ Index .....	90
What is the Index Divisor? .....	90
Example Calculation.....	91
5.9. Exercises on Laspeyres and Paasche Indexes .....	91
Exercise 1: Consumer Goods Basket .....	91
Exercise 2: Stock Market Basket .....	92
Exercise 3: Energy Basket .....	94
Exercise 4: Technology Stock Basket.....	95
Exercise 5: Agricultural Goods Basket .....	96
Summary.....	97
5.10. Proposed Exercises for Laspeyres and Paasche calculation .....	97
Proposed Exercise 1: Technology Gadgets Basket .....	97
Proposed Exercise 2: Food Items Basket .....	98
Proposed Exercise 3: Energy Products Basket .....	98
Proposed Exercise 4: Stock Portfolio Basket.....	99
Proposed Exercise 5: Automobile Basket .....	99
Chapter 6. <b>Two-Dimensional Statistical Analysis</b> .....	101
6.1. Covariance .....	101
Definition.....	101
Formula .....	101
Calculation Steps .....	102
Example.....	102
6.2. Correlation .....	103
Definition.....	103
Formula .....	103
Calculation Steps .....	103
Example.....	103
Interpretation of Correlation Values.....	104
Examples of Correlation Values .....	104
Example.....	105
6.3. Contingency Table .....	106
Example.....	106
Key Concepts.....	106
Application in Business .....	106
6.4. Exercises .....	106
Exercise 1: Customer Satisfaction vs. Purchase Frequency .....	106
Exercise 2: Ad Clicks vs. Device Type.....	107

Exercise 3: Marketing Channel vs. Lead Generation .....	108
Exercise 4: Customer Age vs. Preferred Product Category .....	108
Exercise 5: Email Campaign Type vs. Click-Through Rate .....	109
Chapter 7. <b>Regression Analysis</b> .....	111
7.1. Key Concepts .....	111
7.2. Simple Linear Regression .....	112
Equation .....	112
Interpretation .....	112
7.3. Step-by-step exercises on Regression Calculation .....	112
Exercise 1: Predicting Monthly Sales Based on Advertising Spend ...	112
Exercise 2: Marketing Budget Optimization.....	115
Exercise 3: Predicting Employee Turnover Based on Job Satisfaction	117
Exercise 4: Predicting Product Sales Based on Customer Ratings ...	120
Exercise 5: Predicting Revenue Based on Number of Products Sold ..	122
Exercise 6: Predicting Conversion Rates Based on Email Campaign	
Quality .....	125
Exercise 7: Predicting Website Traffic Based on Marketing Spend....	128
Exercise 8: Predicting Customer Retention Based on Service Quality	131
Exercise 9: Predicting Customer Churn Based on Call Duration .....	134
Exercise 10: Predicting Employee Performance Based on Training	
Hours .....	137
7.4. Proposed Exercises on Regression .....	139
Exercise 1: Predicting Monthly Sales Based on Advertising Spend....	139
Exercise 2: Predicting Customer Satisfaction Based on Service Time	140
Exercise 3: Estimating Product Demand Based on Price .....	140
Exercise 4: Analyzing Employee Performance Based on Training	
Hours .....	140
Exercise 5: Forecasting Revenue Based on Social Media Mentions...	140
Chapter 8. <b>Probability</b> .....	141
8.1. Basic Concepts .....	141
8.2. Conditional Probability .....	141
Definition.....	142
Example.....	142
Interpretation .....	142
8.3. Bayes' Theorem.....	143
Definition.....	143
Example.....	143
Interpretation .....	143
8.4. Applications of Probability and Bayes' Theorem .....	144

Customer Segmentation Using Probability.....	144
Application .....	144
Using Baye’s Theorem.....	144
Interpretation .....	144
8.5. Predicting Customer Churn with Conditional Probability.....	144
Context.....	144
Application .....	145
Using Baye’s Theorem.....	145
Interpretation .....	145
8.6. Optimizing Marketing Campaigns with Multi- Channel Attribution ..	145
Context.....	145
Application .....	145
Solution .....	146
Interpretation .....	146
8.7. Solved Exercises.....	146
Exercise 1: Marketing Campaign Effectiveness.....	146
Exercise 2: Customer Referral Program .....	147
Exercise 3: Product Recall in Manufacturing .....	147
Exercise 4: Email Campaign Sucess Rate .....	148
Exercise 5: Customer Satisfaction Survey.....	148
8.8. Proposed Exercises .....	149
Exercise 1: Customer Referral Program .....	149
Exercise 2: Defective Product in Manufacturing.....	149
Exercise 3: Ad Click Conversion Rate .....	150
Exercise 4: Machine Learning Model Sucess .....	151
Exercise 5: Software Bug Detection .....	151
Chapter 9. <b>Random Variables</b> .....	153
9.1. Discrete Variable .....	153
Binomial Distribution.....	153
Applications in Business and Marketing.....	153
Example: Overbooking in Airline Reservations .....	154
Conclusion .....	155
Poisson Distribution .....	155
Applications in Digital Business and Marketing.....	156
9.2. Binomial Distribution Exercises .....	156
Exercise 1: Email Marketing Campaign.....	156
Exercise 2: Online Ad Click-Through .....	156
Exercise 3: Subscription Service.....	157
Exercise 4: Product Returns .....	157

Exercise 5: Customer Feedback .....	157
9.3. Poisson Distribution Exercises.....	158
Exercise 1: Website Traffic .....	158
Exercise 2: Email Sign-Ups.....	158
Exercise 3: Customer Complaints.....	158
Exercise 4: App Downloads.....	159
Exercise 5: Social Media Mentions .....	159
9.4. Binomial Distribution Proposed Exercises.....	159
Exercise 1: Website Conversion Rate .....	159
Exercise 2: Product Returns .....	160
Exercise 3: Customer Sign-Ups.....	160
Exercise 4: Click-Through Rate .....	160
Exercise 5: Quality Control.....	161
9.5. Poisson Distribution Proposed Exercises .....	161
Exercise 1: Website Visits .....	161
Exercise 2: Customer Service Calls .....	161
Exercise 3: App Installs.....	162
Exercise 4: Online Orders.....	162
Exercise 5: Software Bug Reports.....	162
Chapter 10. <b>Normal Distribution</b> .....	163
10.1. Key Characteristics.....	163
Formula of the Normal Distribution.....	163
10.2. Practical Applications.....	163
10.3. The Standard Normal Distribution .....	164
10.4. Standard normal distribution characteristics .....	164
10.5. Standardizing Any Normal Distribution .....	165
Example 1: Quality Control in Manufacturing.....	165
Example 2: Employee Performance Evaluation .....	165
10.6. Interpreting Z-Scores.....	166
10.7. Using the Standard Normal Distribution Table (Two-Tailed) .....	166
Example 1: Probability of a Value Less Than a Given Value .....	166
Example 2: Probability Between Two Values .....	167
Example 3: Probability Greater Than a Given Value.....	167
10.8. Exercises on Normal Distribution.....	167
Exercise 1: Customer Satisfaction Scores .....	167
Exercise 2: Website Load Times .....	168
Exercise 3: Email Open Rates.....	169
Exercise 4: Product Defect Rates .....	170

Exercise 5: Digital Ad Clicks .....	170
Exercise 6: Sales Performance .....	171
Exercise 7: App Usage Time .....	172
Exercise 8: Return on Investment (ROI) .....	172
10.9. Proposed Exercises on Normal Distribution with Solution .....	173
Exercise 1: Email Campaign Click-Through Rate .....	173
Exercise 2: Customer Retention .....	174
Exercise 3: Social Media Engagement .....	174
Exercise 4: Brand Awareness Survey .....	174
Exercise 5: Advertising Recall .....	175
Exercise 6: Product Return Rate .....	175
Exercise 7: Inventory Turnover .....	175
Exercise 8: Quarterly Sales .....	176
Exercise 9: Employee Productivity .....	176
Exercise 10: Manufacturing Defect Rate .....	176
Exercise 11: Website Bounce Rate .....	177
Exercise 12: App Download Rate .....	177
Exercise 13: Server Downtime .....	177
Exercise 14: Online Transaction Volume .....	178
Exercise 15: Conversion Rate .....	178
Chapter 11. <b>Sampling</b> .....	179
11.1. Central Limit Theorem (CLT) .....	179
Key Points of the Central Limit Theorem .....	179
11.2. Sampling the Mean .....	180
Properties of the Sample Mean .....	180
Practical Example .....	180
11.3. Sampling Proportions .....	180
Properties of the Sample Proportion .....	181
Practical Example .....	181
11.4. Confidence Intervals .....	181
Understanding confidence intervals .....	181
Key Concepts .....	182
When Are Confidence Intervals Used in Real Life? .....	182
Example of Confidence Intervals for the Mean. Estimating Average Monthly Sales .....	183
11.5. Confidence intervals for proportions .....	183
Estimating Customer Satisfaction .....	184
11.6. Proposed Exercises on estimation of confidence intervals with mean sampling .....	184

Exercise 1: Estimating Average Delivery Time.....	184
Exercise 2: Estimating Average Marketing Spend per Campaign....	185
Exercise 3: Estimating Average Page Load Time.....	186
Exercise 4: Estimating Average Customer Lifetime Value (CLV).....	187
Exercise 5: Estimating Average Monthly web Visits .....	188
Exercise 6: Estimating Average Click-Through Rate .....	188
Exercise 7: Estimating the Average Weight of Products.....	189
Exercise 8: Estimating the Average Time Spent on a Website .....	190
Exercise 9: Estimating the Average Monthly Rent .....	191
Exercise 10: Estimating the Average Daily Sales .....	191
Exercise 11: Estimating the Average Commute Time .....	192
Exercise 12: Estimating the Average Lifespan of Batteries.....	193
11.7. Proposed Exercises on estimation of confidence intervals with proportion sampling .....	194
Exercise 1: Estimating Proportion of Satisfied Customers.....	194
Exercise 2: Estimating Proportion of Clicked Ads .....	194
Exercise 3: Estimating Proportion of Repeat Purchasers .....	195
Exercise 4: Estimating Proportion of Defective Products .....	196
Exercise 5: Estimating the Proportion of Returning Customers .....	197
Exercise 6: Estimating the Proportion of Successful Campaigns ....	198
Exercise 7: Estimating the Proportion of High-Engagement Users...	198
Exercise 8: Estimating the Proportion of Products Returned .....	199
Exercise 9: Estimating the Proportion of Positive Reviews.....	200
Exercise 10: Estimating the Proportion of Email Open Rates.....	201
Exercise 11: Estimating the Proportion of Mobile App Conversions...	202
Exercise 12: Estimating the Proportion of Subscribers to a Newsletter	202
Chapter 12. <b>Hypothesis Testing</b> .....	205
12.1. P-Value for Hypothesis Testing .....	206
How to Interpret the p-value? .....	207
12.2. Hypothesis Testing: Examples for Three Scenarios.....	208
Two-Tailed Test Example: Evaluating a New Marketing Strategy ....	208
One-Tailed Test (Upper Tail) Example: Testing a New Feature in a Digital Product .....	209
One-Tailed Test (Lower Tail) Example: Evaluating Cost Reduction in a Business Process .....	210
12.3. Exercises on Hypothesis Testing for the Mean.....	211
Exercise 1: Average Monthly Sales .....	211
Exercise 2: Average Conversion Rate Improvement .....	211
Exercise 3: Average Time on App .....	212
Exercise 4: Reducing Manufacturing Costs.....	213
Exercise 5: Testing Average Customer Satisfaction Score .....	214

Exercise 6: Average Product Price .....	215
Exercise 7: Average Website Load Time .....	215
Exercise 8: Average Daily Sales .....	216
Exercise 9: Average Customer Satisfaction Rating.....	217
Exercise 10: Average Marketing Spend Per Campaign.....	218
12.4. Exercises on Hypothesis Testing for Proportions .....	218
Exercise 1: Proportion of Satisfied Customers .....	218
Exercise 2: Proportion of Returning Customers .....	219
Exercise 3: Proportion of Clicked Ads .....	220
Exercise 4: Proportion of Email Opens .....	221
Exercise 5: Proportion of Positive Reviews .....	222
Exercise 6: Proportion of Successful Campaigns .....	222
Exercise 7: Proportion of Customer Complaints .....	223
Exercise 8: Proportion of Products Returned .....	224
Exercise 9: Proportion of Engaged Users.....	225
Exercise 10: Proportion of Defective Products .....	225
12.5. Proposed Exercises on Hypothesis Testing .....	226
Exercise 1: Average Delivery Time .....	226
Exercise 2: Average Time Spent on Platform .....	226
Exercise 3: Average Number of Items Sold Per Day .....	227
Exercise 4: Average Production Time.....	227
Exercise 5: Email Open Rate.....	227
Exercise 6: Customer Satisfaction .....	228
Exercise 7: User Engagement with New Feature.....	228
Exercise 8: Returned Products .....	228
Exercise 9: Bug Fix Time .....	229
Exercise 10: Consumer preference .....	229
Appendix. <b>Standard Normal Distribution Table</b> .....	230

# 1

# Introduction to Descriptive Statistics

1.1. Key Concepts and Techniques for Frequency Tables. | 1.2. Exercises on Descriptive Statistics. | 1.3. Proposed Exercises.

**D**escriptive statistics are used to summarize and better understand large data sets by providing key metrics such as mean, median, mode, variance, and standard deviation. All of this information is obtained through the use of frequency tables, because they allow you to organize and visualize the raw data into categories or intervals that show the frequency with which each value or range of values occurs. Frequency tables, which allow you to systematically organize and summarize data, are a fundamental tool in data analysis. Frequency tables provide a clear visual representation of how often each value occurs by categorizing and counting the occurrences of certain values within a data set. This approach makes it easier to identify patterns, trends, and anomalies. It helps analysts and decision makers gain insight from raw data.

Frequency tables are used in a variety of fields to understand the distribution of various variables. These tables help quantify the frequency of certain events or responses, whether assessing customer preferences, product performance, or engagement levels. This allows you to make decisions based on evidence instead of guesswork.

In addition, frequency tables are essential for the segmentation of data, the identification of target audiences, and the measurement of the impact of different strategies. Their ability to highlight the most common and least common occurrences makes them an invaluable tool for the refinement of approaches and the optimization of results.

## 1.1. Key Concepts and Techniques for Frequency Tables

### Frequency Table Overview

A **frequency table** is a statistical tool used to represent the distribution of data by grouping values into intervals and calculating various frequencies. Key components include:

- *Absolute Frequency*: The number of times a particular value or range of values occurs.
- *Relative Frequency*: The proportion of occurrences relative to the total number of data points.
- *Cumulative Frequency*: The total of frequencies up to the current point.
- *Cumulative Relative Frequency*: The accumulation of relative frequencies up to the current point.

## Key Calculations

- *Total Frequency*: Sum of all the absolute frequencies.
- *Relative Frequency*: Calculated by dividing the frequency of each category by the total number of observations.

$$\text{Relative Frequency} = \frac{f_i}{N}$$

Where:

- $f_i$  = Frequency of a category
- $N$  = Total number of observations
- *Cumulative Absolute Frequency*: The running total of absolute frequencies up to a certain class interval.
- *Cumulative Relative Frequency*: Accumulation of relative frequencies up to the current class interval, providing insight into the percentage of data that falls within a range.

$$\text{Cumulative Relative Frequency} = \sum \frac{f_i}{N}$$

## Example Step-by-Step Frequency Table Calculation

Let's consider the following data set representing exam scores from 16 students:

5, 7, 8, 7, 6, 9, 5, 6, 7, 8, 9, 5, 6, 9, 10, 7, 7

- *Step 1: Organize the data.* Sort the data in ascending order:

5, 5, 5, 6, 6, 6, 7, 7, 7, 7, 7, 8, 8, 9, 9, 10

- *Step 2: Count the frequencies.* Now count how many times each value appears:
  - The value 5 appears 3 times.
  - The value 6 appears 3 times.
  - The value 7 appears 5 times.
  - The value 8 appears 2 times.
  - The value 9 appears 2 times.
  - The value 10 appears 1 time.
- *Step 3: Create the frequency table.*

Value	Absolute Frequency
5	3
6	3
7	5
8	2
9	2
10	1

- *Step 4: Calculate relative frequency.* To calculate the relative frequency, divide the absolute frequency of each value by the total number of data points (16):

$$\text{Relative Frequency} = \frac{\text{Absolute Frequency}}{16}$$

Value	Absolute Frequency	Relative Frequency
5	3	0.1875
6	3	0.1875
7	5	0.3125
8	2	0.1250
9	2	0.1250
10	1	0.0625

- *Step 5: Calculate cumulative frequencies.* Finally, calculate the cumulative absolute and cumulative relative frequencies:

Value	Absolute Frequency	Relative Frequency	Cumulative Absolute Frequency	Cumulative Relative Frequency
5	3	0.1875	3	0.1875
6	3	0.1875	6	0.3750
7	5	0.3125	11	0.6875
8	2	0.1250	13	0.8125
9	2	0.1250	15	0.9375
10	1	0.0625	16	1.0000

### *Example 1: Statistical Analysis of Family Units in a Housing Development*

In a housing development, a study has been conducted on the number of members that make up the family unit, yielding the following results:

Data:

- Number of People: 2, 3, 4, 5, 6, 7
- Number of Families: 3, 10, 16, 15, 9, 7

### *Example 2: Statistical Analysis of Family Units in a Neighborhood*

In a neighborhood, a study has been conducted to determine the number of members that make up the family unit, and the results are presented in the adjacent table:

Data:

- Number of People: 2, 3, 4, 5, 6, 7
- Number of Families: 28, 63, 99, 112, 42, 0

### *Explanation of Variables*

- *Number of People:* This is the variable in the study because it represents the number of members in each family unit. A variable is something that can change, and in this case, the number of people per family can vary.

- *Number of Families:* This represents the frequency or how often a family with a certain number of members is observed in the study. Frequency indicates the count of occurrences for each value of the variable.

### Why Order Matters

The data is initially unordered in terms of processing, but it is important to organize it to understand how families are distributed across different family sizes. Ordering the data makes it easier to calculate cumulative frequencies and relative frequencies, which are essential for understanding the distribution.

## Step-by-Step Solution

### Example 1 (Housing Development)

Ordered Data:

Number of People	Number of Families
2	3
3	10
4	16
5	15
6	9
7	7

- *Step 1: Total Frequency.* The total frequency is the sum of all the families:

$$N = 3 + 10 + 16 + 15 + 9 + 7 = 60$$

- *Step 2: Relative Frequency.* The relative frequency for each category is calculated by dividing the number of families by the total frequency.

$$f_i = \frac{n_i}{N}$$

For example, for “Number of People = 2”:

$$f_1 = \frac{3}{60} = 0.05$$

Repeating for all categories:

Number of People	Number of Families	Relative Frequency
2	3	0.05
3	10	0.16
4	16	0.26
5	15	0.25
6	9	0.15
7	7	0.11

- *Step 3: Cumulative Absolute Frequency.* The cumulative absolute frequency is calculated by adding each frequency to the sum of the previous frequencies:

Number of People	Cumulative Absolute Frequency
2	3
3	3 + 10 = 13
4	13 + 16 = 29
5	29 + 15 = 44
6	44 + 9 = 53
7	53 + 7 = 60

- *Step 4: Cumulative Relative Frequency.* The cumulative relative frequency is calculated by adding each relative frequency to the sum of the previous relative frequencies:

Number of People	Cumulative Relative Frequency
2	0.05
3	0.05 + 0.167 = 0.217
4	0.217 + 0.267 = 0.484
5	0.484 + 0.25 = 0.734
6	0.734 + 0.15 = 0.884
7	0.884 + 0.117 = 1.000

### Example 2 (Neighborhood)

Ordered Data:

Number of People	Number of Families
2	28
3	63
4	99
5	112
6	42
7	0

- *Step 1: Total Frequency.* The total frequency is the sum of all the families:

$$N = 28 + 63 + 99 + 112 + 42 + 0 = 344$$

- *Step 2: Relative Frequency.* The relative frequency for each category is calculated by dividing the number of families by the total frequency. For example, for “Number of People = 2”:

$$f_1 = \frac{28}{344} \approx 0.0814$$

Repeating for all categories:

Number of People	Number of Families	Relative Frequency
2	28	0.0814
3	63	0.1831
4	99	0.2872
5	112	0.3256
6	42	0.1221
7	0	0.0000

- *Step 3: Cumulative Absolute Frequency.* The cumulative absolute frequency is calculated as follows:

Number of People	Cumulative Absolute Frequency
2	28
3	$28 + 63 = 91$
4	$91 + 99 = 190$
5	$190 + 112 = 302$
6	$302 + 42 = 344$
7	344

- *Step 4: Cumulative Relative Frequency.* The cumulative relative frequency is calculated as follows:

Number of People	Cumulative Relative Frequency
2	0.0814
3	$0.0814 + 0.1831 = 0.2645$
4	$0.2645 + 0.2872 = 0.5517$
5	$0.5517 + 0.3256 = 0.8773$
6	$0.8773 + 0.1221 = 0.9994$
7	1.0000

### Comparison of Results

- In the housing development, the total number of families was 60, while in the neighborhood, it was 344.
- The distribution of families is more concentrated around 4 and 5 people in both cases, but the neighborhood had a larger number of families overall.
- The neighborhood shows a higher frequency for larger families (5 people) compared to the housing development, where 4 people families are more common.

## 1.2. Exercises on Descriptive Statistics

### Exercise 1: Daily Visits to a Website

Analyze the distribution of daily visits to a website using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

#### Data Table

Visits per Day	$n_i$ (Absolute Frequency)
0-5	35
5-10	50
10-15	60
15-20	45
20-25	40

### Step-by-Step Calculations

To fully understand the distribution of daily visits to a marketing website, we need to calculate the following:

1. *Total Frequency*. The sum of all the frequencies. The total frequency is calculated by summing up all the frequencies:

$$N = 35 + 50 + 60 + 45 + 40 = 230$$

2. *Relative Frequency*. The percentage of each frequency relative to the total frequency. The relative frequency for each category is calculated by dividing the frequency of each category by the total frequency:

$$f_i = \frac{n_i}{N}$$

For example, for the category “0-5”:

$$f_1 = \frac{35}{230} \approx 0.1522$$

3. *Cumulative Absolute Frequency*. The running total of the frequencies. The cumulative absolute frequency is calculated by adding each frequency to the sum of the frequencies that came before it. For example:

- For “0-5”: 35
- For “5-10”:  $35 + 50 = 85$
- For “10-15”:  $85 + 60 = 145$
- For “15-20”:  $145 + 45 = 190$
- For “20-25”:  $190 + 40 = 230$

4. *Cumulative Relative Frequency*. The running total of the relative frequencies. The cumulative relative frequency is calculated by adding each relative frequency to the sum of the relative frequencies that came before it. For example:

- For “0-5”: 0.1522
- For “5-10”:  $0.1522 + 0.2174 = 0.3696$
- For “10-15”:  $0.3696 + 0.2609 = 0.6305$
- For “15-20”:  $0.6305 + 0.1957 = 0.8261$
- For “20-25”:  $0.8261 + 0.1739 = 1.0000$

### Final Calculations

Visits per Day	$n_i$ (Absolute Frequency)	$f_i$ (Relative Frequency)	$N_i$ (Cumulative Absolute Frequency)	$F_i$ (Cumulative Relative Frequency)
0-5	35	0.1522	35	0.1522
5-10	50	0.2174	85	0.3696
10-15	60	0.2609	145	0.6305
15-20	45	0.1957	190	0.8261
20-25	40	0.1739	230	1.0000

## Exercise 2: Number of Products Sold per Day (Marketing)

Analyze the distribution of the number of products sold per day using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

### Data Table

Products Sold per Day	$n_i$ (Absolute Frequency)
0-10	40
10-20	55
20-30	65
30-40	50
40-50	45

### Step-by-Step Calculations

To fully understand the distribution of products sold per day, we need to calculate the following:

1. *Total Frequency*. The sum of all the frequencies. The total frequency is calculated by summing up all the frequencies:

$$N = 40 + 55 + 65 + 50 + 45 = 255$$

2. *Relative Frequency*. The percentage of each frequency relative to the total frequency. The relative frequency for each category is calculated by dividing the frequency of each category by the total frequency:

$$f_i = \frac{n_i}{N}$$

For example, for the category “0-10”:

$$f_1 = \frac{40}{255} \approx 0.1569$$

3. *Cumulative Absolute Frequency*. The running total of the frequencies. The cumulative absolute frequency is calculated by adding each frequency to the sum of the frequencies that came before it. For example:

- For “0-10”: 40
- For “10-20”:  $40 + 55 = 95$
- For “20-30”:  $95 + 65 = 160$
- For “30-40”:  $160 + 50 = 210$
- For “40-50”:  $210 + 45 = 255$

4. *Cumulative Relative Frequency*. The running total of the relative frequencies. The cumulative relative frequency is calculated by adding each relative frequency to the sum of the relative frequencies that came before it. For example:

- For “0-10”: 0.1569
- For “10-20”:  $0.1569 + 0.2157 = 0.3725$
- For “20-30”:  $0.3725 + 0.2549 = 0.6275$
- For “30-40”:  $0.6275 + 0.1961 = 0.8235$
- For “40-50”:  $0.8235 + 0.1765 = 1.0000$

### Final Calculations

Products Sold per Day	$n_i$ (Absolute Frequency)	$f_i$ (Relative Frequency)	$N_i$ (Cumulative Absolute Frequency)	$F_i$ (Cumulative Relative Frequency)
0-10	40	0.1569	40	0.1569
10-20	55	0.2157	95	0.3725
20-30	65	0.2549	160	0.6275
30-40	50	0.1961	210	0.8235
40-50	45	0.1765	255	1.0000

## Exercise 3: Customer Reviews per Day (Marketing)

Analyze the distribution of customer reviews per day using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

### Data Table

Reviews per Day	$n_i$ (Absolute Frequency)
0-2	50
2-4	60
4-6	66
6-8	54
8-10	30

### Step-by-Step Calculations

To fully understand the distribution of customer reviews per day, we need to calculate the following:

1. *Total Frequency*: The sum of all the frequencies. The total frequency is calculated by summing up all the frequencies:

$$N = 50 + 60 + 66 + 54 + 30 = 260$$

2. *Relative Frequency*: The percentage of each frequency relative to the total frequency. The relative frequency for each category is calculated by dividing the frequency of each category by the total frequency:

$$f_i = \frac{n_i}{N}$$

For example, for the category “0-2”:

$$f_1 = \frac{50}{260} \approx 0.1923$$

3. *Cumulative Absolute Frequency*: The running total of the frequencies. The cumulative absolute frequency is calculated by adding each frequency to the sum of the frequencies that came before it. For example:
  - For “0-2”: 50
  - For “2-4”:  $50 + 60 = 110$
  - For “4-6”:  $110 + 66 = 176$
  - For “6-8”:  $176 + 54 = 230$
  - For “8-10”:  $230 + 30 = 260$
4. *Cumulative Relative Frequency*: The running total of the relative frequencies. The cumulative relative frequency is calculated by adding each relative frequency to the sum of the relative frequencies that came before it. For example:
  - For “0-2”: 0.1923
  - For “2-4”:  $0.1923 + 0.2308 = 0.4231$
  - For “4-6”:  $0.4231 + 0.2538 = 0.6769$
  - For “6-8”:  $0.6769 + 0.2077 = 0.8846$
  - For “8-10”:  $0.8846 + 0.1154 = 1.0000$

### Final Calculations

Reviews per Day	$n_i$ (Absolute Frequency)	$f_i$ (Relative Frequency)	$N_i$ (Cumulative Absolute Frequency)	$F_i$ (Cumulative Relative Frequency)
0-2	50	0.1923	50	0.1923
2-4	60	0.2308	110	0.4231
4-6	66	0.2538	176	0.6769
6-8	54	0.2077	230	0.8846
8-10	30	0.1154	260	1.0000

## Exercise 4: Daily Transactions in a Business

Analyze the distribution of daily transactions in a business using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

### Data Table

Transactions per Day	$n_i$ (Absolute Frequency)
0-20	40
20-40	60
40-60	55
60-80	50
80-100	45

### Step-by-Step Calculations

To fully understand the distribution of daily transactions in a business, we need to calculate the following:

1. *Total Frequency*: The sum of all the frequencies. The total frequency is calculated by summing up all the frequencies:

$$N = 40 + 60 + 55 + 50 + 45 = 250$$

2. *Relative Frequency*: The percentage of each frequency relative to the total frequency. The relative frequency for each category is calculated by dividing the frequency of each category by the total frequency:

$$f_i = \frac{n_i}{N}$$

For example, for the category “0-20”:

$$f_1 = \frac{40}{250} = 0.1600$$

3. *Cumulative Absolute Frequency*: The running total of the frequencies. The cumulative absolute frequency is calculated by adding each frequency to the sum of the frequencies that came before it. For example:
  - For “0-20”: 40
  - For “20-40”:  $40 + 60 = 100$
  - For “40-60”:  $100 + 55 = 155$
  - For “60-80”:  $155 + 50 = 205$
  - For “80-100”:  $205 + 45 = 250$
4. *Cumulative Relative Frequency*: The running total of the relative frequencies. The cumulative relative frequency is calculated by adding each relative frequency to the sum of the relative frequencies that came before it. For example:
  - For “0-20”: 0.1600
  - For “20-40”:  $0.1600 + 0.2400 = 0.4000$
  - For “40-60”:  $0.4000 + 0.2200 = 0.6200$
  - For “60-80”:  $0.6200 + 0.2000 = 0.8200$
  - For “80-100”:  $0.8200 + 0.1800 = 1.0000$

### Final Calculations

Transactions per Day	$n_i$ (Absolute Frequency)	$f_i$ (Relative Frequency)	$N_i$ (Cumulative Absolute Frequency)	$F_i$ (Cumulative Relative Frequency)
0-20	40	0.1600	40	0.1600
20-40	60	0.2400	100	0.4000
40-60	55	0.2200	155	0.6200
60-80	50	0.2000	205	0.8200
80-100	45	0.1800	250	1.0000

## Exercise 5: Units Produced per Day (Business)

Analyze the distribution of units produced per day using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

### Data Table

Units Produced per Day	$n_i$ (Absolute Frequency)
0-10	35
10-20	50
20-30	60
30-40	55
40-50	40

### Step-by-Step Calculations

To fully understand the distribution of units produced per day, we need to calculate the following

1. *Total Frequency*: The sum of all the frequencies. The total frequency is calculated by summing up all the frequencies:

$$N = 35 + 50 + 60 + 55 + 40 = 240$$

2. *Relative Frequency*: The percentage of each frequency relative to the total frequency. The relative frequency for each category is calculated by dividing the frequency of each category by the total frequency:

$$f_i = \frac{n_i}{N}$$

For example, for the category “0-10”:

$$f_1 = \frac{35}{240} \approx 0.1458$$

3. *Cumulative Absolute Frequency*: The running total of the frequencies. The cumulative absolute frequency is calculated by adding each frequency to the sum of the frequencies that came before it. For example:

- For “0-10”: 35
- For “10-20”:  $35 + 50 = 85$
- For “20-30”:  $85 + 60 = 145$
- For “30-40”:  $145 + 55 = 200$
- For “40-50”:  $200 + 40 = 240$

4. *Cumulative Relative Frequency*: The running total of the relative frequencies. The cumulative relative frequency is calculated by adding each relative frequency to the sum of the relative frequencies that came before it. For example:

- For “0-10”: 0.1458
- For “10-20”:  $0.1458 + 0.2083 = 0.3542$
- For “20-30”:  $0.3542 + 0.2500 = 0.6042$
- For “30-40”:  $0.6042 + 0.2292 = 0.8333$
- For “40-50”:  $0.8333 + 0.1667 = 1.0000$

### Final Calculations

Units Produced per Day	$n_i$ (Absolute Frequency)	$f_i$ (Relative Frequency)	$N_i$ (Cumulative Absolute Frequency)	$F_i$ (Cumulative Relative Frequency)
0-10	35	0.1458	35	0.1458
10-20	50	0.2083	85	0.3542
20-30	60	0.2500	145	0.6042
30-40	55	0.2292	200	0.8333
40-50	40	0.1667	240	1.0000

## 1.3. Proposed Exercises

### Exercise 1: Daily Visits to an E-commerce Website

Analyze the distribution of daily visits to an e-commerce website using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

#### Data Table

Visits per Day	Frequency
0-10	45
10-20	55
20-30	50
30-40	40
40-50	35
50-60	25

#### Task

1. Calculate the total frequency.
2. Calculate the relative frequency for each category.
3. Calculate the cumulative absolute frequency.
4. Calculate the cumulative relative frequency.

## Exercise 2: Number of Products Sold per Week (Marketing)

Analyze the distribution of the number of products sold per week using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

### Data Table

Products Sold per Week	Frequency
0-50	30
50-100	45
100-150	55
150-200	40
200-250	35
250-300	25

### Task

1. Calculate the total frequency.
2. Calculate the relative frequency for each category.
3. Calculate the cumulative absolute frequency.
4. Calculate the cumulative relative frequency.

## Exercise 3: Customer Support Tickets per Day (Digital Business)

Analyze the distribution of customer support tickets received per day using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

### Data Table

Tickets per Day	Frequency
0-10	40
10-20	50
20-30	55
30-40	45
40-50	35
50-60	30

### Task

1. Calculate the total frequency.
2. Calculate the relative frequency for each category.
3. Calculate the cumulative absolute frequency.
4. Calculate the cumulative relative frequency.

## Exercise 4: Monthly Subscriptions to a Digital Service

Analyze the distribution of monthly subscriptions to a digital service using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

### Data Table

Subscriptions per Month	Frequency
0-50	35
50-100	45
100-150	50
150-200	40
200-250	30
250-300	20

### Task

1. Calculate the total frequency.
2. Calculate the relative frequency for each category.
3. Calculate the cumulative absolute frequency.
4. Calculate the cumulative relative frequency.

## Exercise 5: Daily Transactions in an Online Store

Analyze the distribution of daily transactions in an online store using the provided data. Calculate the total frequency, relative frequency, cumulative absolute frequency, and cumulative relative frequency.

### Data Table

Transactions per Day	Frequency
0-20	35
20-40	50
40-60	55
60-80	45
80-100	40
100-120	30

### Task

1. Calculate the total frequency.
2. Calculate the relative frequency for each category.
3. Calculate the cumulative absolute frequency.
4. Calculate the cumulative relative frequency.