ALRIC INFO TECH

(A Software & Analytics Company in Trivandrum)

AN INTERN REPORT

Submitted By,

SREE KRISHNA NIVETHITHA.S

960223243116

in partial fulfilment for the award of the degree

of

BACHELOR OF TECHNOLOGY

IN

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

ARUNACHALA COLLEGE OF ENGINEERING FOR WOMEN

JULY 2025

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to the management of Arunachala College of Engineering for Women for providing me with the opportunity to pursue this internship programme as a part of my academic curriculum. I extend my sincere thanks to our respected Chairman, **Dr. T. Krishnaswamy, M.E., Ph.D.,** whose vision and encouragement have always motivated students to achieve excellence.

I am equally grateful to our Principal, **Dr. S. Joseph Jawahar, M.E., Ph.D., MBA, M.I.E.,** for his constant support, guidance, and valuable suggestions, which have greatly contributed to my academic growth.

My heartfelt thanks go to **Dr. T. Sunitha, M.E., Ph.D.,** Head of the Department of Artificial Intelligence and Data Science, for her continuous encouragement, valuable advice, and motivation throughout my journey. Her guidance has been instrumental in shaping my interest in the field and has given me the confidence to take up this internship with enthusiasm.

I sincerely thank **Alric Info Tech, Trivandrum**, for offering me this opportunity to undergo a one-month internship programme. The training sessions and projects provided me with practical exposure that bridged the gap between theoretical knowledge and industry applications. I am deeply thankful to my trainers and mentors at the company for their constant guidance, knowledge sharing, and support in helping me understand real-time problem-solving techniques.

This internship has been a valuable milestone in my academic journey, and I sincerely acknowledge every individual and institution that has contributed to making this experience a meaningful and successful one.

CERTIFICATE

ALRIC INFOTECH Pvt Ltd
Ruling The Technology

Date: 24.07.2025

INTERNSHIP CERTIFICATE

This is to certify that **Ms. SREE KRISHNA NIVETHITHA S, Reg. No (960223243116)**, Student of Dept. of Artificial Intelligence and Data Science, **ARUNACHALA COLLEGE OF ENGINEERING FOR WOMEN, NAGERCOIL** has successfully completed Internship in Data Analytics at Alric Infotech Pvt Ltd, Trivandrum starting from 23.06.2025 to 23.07.2025.

During the period, she has successfully completed a project and also she was methodical and hardworking. She shows a very good analytical and logical approach while studying and developing systems, along with remarkable communication skills.

We wish **Ms. SREE KRISHNA NIVETHITHA S** all the best for all the future endeavours.

For **Alric Infotech Pvt Ltd**

Arun Alfred

Managing Director

Authorized Signatory

+91 8089421289

3rd Floor, Kairali Complex, GPO Lane,
Mele Thampanoor, Thiruvananthapuram
695001, Kerala, India

www.alricinfotech.com
info@alricinfotech.com

CIN: U72900KL2022PTC077617

INTRODUCTION

The internship programme forms a vital part of academic learning as it allows students to step outside the classroom and gain practical exposure to real-world working environments. It is designed not only to provide hands-on experience but also to create a bridge between theoretical concepts and industry-level practices. As part of my curriculum, I successfully completed a 30-day internship at Alric Info Tech, Trivandrum, located at the 3rd Floor, Kairali Complex, GPO Lane, Mele Thampanoor, Kerala, from 23rd June 2025 to 23rd July 2025. This opportunity allowed me to immerse myself in the professional world of data analytics, where I was able to apply my academic knowledge to practical tasks and projects.

During this internship, I was introduced to various aspects of Data handling, Analysis, Visualization, and Automation, which are fundamental components of modern data-driven decision-making. I worked with tools that are widely used in the industry, such as Microsoft Excel, Power BI, and Python libraries including Pandas, NumPy, Matplotlib, and Seaborn. These tools not only enhanced my technical proficiency but also helped me understand how raw data can be transformed into meaningful insights for business improvement.

In addition to technical learning, the internship provided me with valuable lessons in Soft skills development. Working under the guidance of experienced professionals, I learned the importance of time management, effective communication, and collaboration in a team setting. I was also encouraged to approach problem-solving with creativity and logical thinking, which gave me the confidence to tackle real-time challenges.

OBJECTIVES OF INTERNSHIP

- Gain practical experience in data analytics tools by applying theoretical knowledge to industry practices and learning how business data is collected, cleaned, and transformed into insights.
- Understand the complete workflow of business data — collection, storage, analysis, visualization, and reporting — to support business intelligence.
- Enhance Excel skills using Power Query, Pivot Tables, and Dashboards for efficient data handling and reporting.
- Apply statistical techniques like probability, sampling, distribution fitting, and hypothesis testing to solve real-world data problems.
- Learn Power BI to create interactive dashboards and reports, improving data visualization for stakeholders.
- Use Python libraries such as Pandas, NumPy, Matplotlib, and Seaborn for advanced analysis, visualization, and predictive modeling.
- Improve problem-solving and critical thinking skills by working on real datasets and deriving actionable recommendations.
- Build communication and presentation skills by effectively conveying data-driven insights to diverse audiences.
- Collaborate with team members and industry experts to gain exposure to professional work environments and enhance teamwork skills.
- Adapt to new technologies and methodologies, developing resilience and flexibility to meet evolving industry demands.

ORGANIZATION PROFILE

Alric Info Tech is a growing technology solutions company offering a wide range of IT services, training, and consultancy. The organization specializes in software development, analytics, and business intelligence, providing innovative solutions that enhance business efficiency and decision-making.

Alongside its services, the company has a strong training division focusing on emerging technologies such as Artificial Intelligence, Data Science, and Data Analytics. Through workshops, practical sessions, and project-based learning, it equips students and professionals with industry-ready skills.

The internship program at Alric Info Tech is designed to bridge the gap between academics and industry. Guided by experienced professionals, interns engage in real-time projects and problem-solving tasks, gaining both technical expertise and hands-on experience.

By integrating services, training, and practical exposure, Alric Info Tech has established itself as both a trusted technology partner for businesses and a skill-development platform for learners.

Under the mentorship of Arun Alfred, the internship provides valuable guidance and coaching, helping interns apply their skills effectively and enhance their professional growth.

The organization promotes a supportive and collaborative environment where interns are encouraged to explore new technologies, solve challenges, and build confidence for their future careers.

INTERNSHIP WORK AND ACTIVITIES

During my 30-day internship at Alric Info Tech, I had the opportunity to explore a wide range of tools, technologies, and methodologies that are essential in the field of Data Analytics and Business Intelligence (BI). The internship was carefully structured so that each week I learned new concepts, beginning with foundational data handling techniques using Excel, moving through statistical methods, and advancing towards real-world applications with tools like Power BI and Python.

This report details the activities I performed, the tools I learned, and how they are applied in solving practical data problems, while also highlighting the importance of analytical thinking and data-driven decision-making in the industry.

Excel Analytics – Data Handling & Visualization

- Microsoft Excel formed the cornerstone of my initial training. While often perceived as a basic tool, Excel's advanced features offer powerful capabilities for cleaning, transforming, and visualizing data. My training focused on using Excel to efficiently process datasets, extract meaningful insights, and present data in an interactive and interpretable manner.

- Used Power Query to import and process large datasets from CSV files, text files, and online sources.

- Performed data cleaning tasks including removing duplicates, handling missing values, and standardizing inconsistent entries.

- Example: Worked on a customer sales dataset, corrected spelling errors, merged tables containing customer and transaction details, and transformed raw data into structured formats ready for analysis.

- Explored formulas such as VLOOKUP, INDEX, MATCH, and logical

functions like IF, AND, and OR to manipulate and retrieve data efficiently.

- Used Pivot Tables to summarize large datasets and analyze trends, such as total sales by region, product category, and monthly growth.
- Applied conditional formatting to highlight key insights, for example, sales below targets in red and top-performing regions in green.
- Designed interactive dashboards combining charts, slicers, and Pivot Tables, enabling users to filter data dynamically and explore key metrics in real time.
- Created a sales dashboard where users could filter by year, location, or product category to analyze revenue trends, sales growth, and product comparisons.
- Utilized charts such as line graphs, bar charts, and pie charts to communicate insights clearly and effectively.

Statistics &Probability – Strengthening Analytical Thinking

- Statistical methods were a critical part of the internship, helping me adopt a structured, data-driven approach to problem-solving. I applied these techniques to understand patterns, forecast outcomes.
- Applied concepts such as permutations, combinations, and probability distributions to explore data patterns and predict outcomes.
- Used sampling methods to analyze large datasets efficiently by examining representative subsets rather than processing all records.
- Applied probability distributions to calculate customer behavior predictions, such as churn rates, which helped organizations assess risks and opportunities.
- Conducted hypothesis testing to validate assumptions and support data-driven decisions.

- Applied t-tests and chi-square tests to assess whether new initiatives, such as product launches, significantly impacted performance metrics.
- Learned the importance of making business decisions based on statistical evidence rather than assumptions**.**

Business Intelligence Tools – Applying Analytics Platforms

- In the latter part of the internship, I gained hands-on experience with modern Business Intelligence tools, which are widely used in the industry to extract insights from complex datasets and make informed decisions.
- Worked with Power BI to integrate multiple datasets and build robust data models.
- Used Power Query in Power BI for data cleaning and transformation.
- Created calculated fields and advanced measures using DAX (Data Analysis Expressions), including cumulative totals and growth metrics.
- Designed a sales performance dashboard integrating sales, regions, and product data, providing interactive filters and visualizations to explore trends and key performance indicators.
- Published dashboards on Power BI Service (PowerBI.com) to enable real-time collaboration and sharing across teams.
- Worked with Python for data analytics to handle larger datasets and perform complex computations.
- Set up Jupyter Notebook as an interactive environment for writing, testing, and visualizing data analyses.

- Used Pandas for data manipulation, including importing datasets, cleaning missing values, grouping data, and generating summaries.
- Applied NumPy for fast numerical computations, such as averages, standard deviations, and array operations.
- Created visualizations using Matplotlib and Seaborn, including bar charts, histograms, and heatmaps, to make complex datasets interpretable and actionable.
- Leveraged Python to explore patterns, make predictions, and communicate insights effectively through visual storytelling.
- Collaborated with team members to solve real-world data challenges, applying analytical techniques to generate actionable insights that supported decision-making processes.

SKILLS LEARNED

During the course of my internship, I was able to acquire a blend of technical skills and soft skills. These skills not only enhanced my knowledge of data analytics but also shaped my ability to work effectively in a professional environment.

TECHNICAL SKILLS

1. Data Cleaning, Preparation, and Transformation

- I learned how to handle raw and unstructured data by applying various cleaning techniques such as handling missing values, removing duplicates, and standardizing data formats.
- Tools like Excel Power Query, Python's Pandas, and Power BI helped me transform datasets into a structured and usable format.

2. **Data Visualization using Excel, Power BI**

- I gained hands-on experience in presenting complex data through meaningful visualizations.
- Excel allowed me to create quick dashboards and charts for day-to-day analysis, while Power BI provided advanced visualization techniques such as heatmaps, trend lines, and interactive dashboards.
- These skills taught me how visual storytelling makes data more understandable to non-technical audiences and decision-makers.
- Learned how visual storytelling helps make complex datasets more understandable and supports informed decision-making.
- Developed the ability to design user-friendly dashboards that allow stakeholders to explore data dynamically and gain actionable insights.
- Enhanced skills in selecting appropriate chart types and formatting techniques to highlight key trends and comparisons clearly.

3. **Python for Analytics (Pandas, NumPy, Matplotlib, Seaborn)**

- Used Pandas for manipulating datasets, cleaning data, handling missing values, and generating summaries.
- Applied NumPy for numerical operations, including averages, standard deviations, and array-based calculations.
- Utilized Matplotlib and Seaborn for creating visualizations such as bar charts, histograms, and heatmaps, making complex data easier to interpret.
- Implemented data aggregation and grouping techniques to extract meaningful insights from large datasets efficiently.

- Applied basic predictive analytics, using Python to identify trends, correlations, and patterns that support data-driven decision-making.

5.Statistical Analysis and Hypothesis Testing

- Applied sampling techniques to efficiently analyze large datasets by examining representative subsets instead of all records.
- Used probability distributions to forecast outcomes, predict customer behavior, and identify potential risks or opportunities.
- Conducted hypothesis testing using t-tests and chi-square tests to validate assumptions, for example, evaluating whether a new business strategy significantly improved sales.

SOFT SKILLS

1.Problem-Solving and Analytical Thinking

- I learned to approach data problems systematically—identifying the issue, exploring the dataset, applying tools, and finally interpreting results.
- This structured problem-solving mindset is essential in both academic research and industry practice.
- Gained experience in drawing actionable insights from complex datasets, enabling me to propose data-driven recommendations to support decision-making.

2.Team Collaboration and Communication

- Throughout the internship, I interacted with trainers, mentors, and fellow interns.

- I realized that effective communication is as important as technical knowledge, especially when presenting results to stakeholders who may not have a technical background.
- Learned to collaborate on group projects and share insights effectively, ensuring alignment and contributing to successful team outcomes.

3.Time Management and Project Handling

- Completing multiple tasks within the given 30-day period required careful planning and prioritization.
- This experience taught me to manage time effectively and complete tasks .

4.Report Preparation and Presentation Skills

- Preparing this detailed internship report enhanced my ability to structure information logically and present it in a professional format.
- It also improved my confidence in documenting technical findings and sharing them with an audience.
- Learned to summarize complex data and analyses concisely, making insights easier for stakeholders to understand.
- Developed skills in visual presentation of information, including charts, tables, and dashboards, to support clear and impactful reporting.

5. Adaptability and Learning Agility

- During the internship, I had to quickly learn new tools, techniques, and workflows to meet project requirements.
- This adaptability improved my ability to adjust to new environments.

- Gained experience in switching between multiple tools and platforms such as Excel, Power BI, and Python, depending on the task requirements.
- Developed the ability to pick up new analytical concepts and techniques rapidly, which helped in completing projects efficiently.
- Learned to respond positively to feedback from mentors and peers, using it to refine my work and improve outcomes.
- Strengthened my self-learning and research skills, enabling me to explore solutions independently and apply them effectively in practical scenarios.

6.Leadership and Initiative

- Took the initiative to independently research and apply new methods when faced with unfamiliar challenges.
- Learned to take ownership of tasks, ensuring they were completed on time and met quality standards.
- Developed leadership qualities by guiding peers during collaborative activities, offering suggestions, and encouraging knowledge sharing.
- Built confidence in taking proactive steps, rather than waiting for instructions, which enhanced productivity and problem ownership.
- Gained confidence in making informed decisions under time constraints.
- Strengthened the ability to think logically and critically, which helped in identifying gaps, avoiding errors, and delivering accurate results.

INTRODUCTION ABOUT THE PROJECT

This project focuses on analyzing the Blinkit grocery dataset to study sales trends, customer preferences, and outlet performance. The data was cleaned using Python in VS Code and visualized through an interactive Power BI dashboard to generate meaningful business insights.

PROJECT DESCRIPTION

The project focuses on analyzing the Blinkit grocery dataset to gain valuable insights into consumer behavior, sales trends, product performance, and operational efficiency. With the rising demand for online grocery delivery, businesses need to leverage data-driven insights to optimize inventory, improve customer satisfaction, and maximize revenue. This project explores Blinkit's transactional data to identify key patterns, challenges, and opportunities for improvement. The findings from this analysis can support strategic decision-making in sales, marketing, and supply chain management.

OBJECTIVES OF THE PROJECT

The major objectives of the project are as follows:

1. To analyze order trends and identify peak sales periods.
2. To study product performance, including best-selling and low-selling items.
3. To understand customer purchase behavior and preferences.
4. To evaluate delivery efficiency and order fulfillment patterns.
5. To identify operational bottlenecks and suggest areas for improvement in grocery delivery.

6. To analyze regional and category-wise sales patterns for better inventory planning.
7. To provide actionable insights and recommendations for optimizing Blinkit's overall business operations.

SCOPE OF THE PROJECT

- **Sales and Order Analysis:** Study of historical sales data to identify patterns, seasonal variations, and demand fluctuations.
- **Customer Insights:** Analysis of customer buying habits to enable targeted promotions and personalized marketing strategies.
- **Product Performance:** Evaluation of different product categories to inform stock management and inventory optimization.
- **Operational Efficiency:** Visualization and analysis of delivery timelines and order fulfillment data to enhance supply chain effectiveness.
- **Challenges and Recommendations:** Identification of gaps in sales, logistics, and operations, along with data-driven recommendations for improvement.

METHODOLOGY

The methodology adopted for this project involves a structured process of data collection, cleaning, transformation, visualization, and analysis to extract meaningful insights from the Blinkit dataset. The steps are outlined below:

1.Dataset Collection

- The Blinkit grocery dataset was obtained from an open-source repository.

- The dataset contained information on product details, outlet types, sales, ratings, and customer preferences.

2.Data Cleaning (in VS Code using Python and Pandas)

- The raw dataset was imported into VS Code and processed using the Pandas library.
- Missing values were handled by filling with mean, median, or mode where appropriate.
- Duplicate records were removed to ensure data consistency.
- Text columns such as Item Fat Content were standardized for uniformity.
- Outliers were detected and treated to improve analysis reliability.
- The cleaned dataset was then saved as a new file for visualization.

STEPS TO CLEAN DATASET IN VS CODE

1.Import Required Libraries

```
import pandas as pd
```

✓ 1.7s Python

2.Load the Dataset

```
df = pd.read_excel(r'C:\Users\HP\Desktop\intern report\BlinkIT Grocery Data (2).xlsx')
df
```

Python

Output:

	Item Fat Content	Item Identifier	Item Type	Outlet Establishment Year	Outlet Identifier	Outlet Location Type	Outlet Size	Outlet Type	Item Visibility	Item Weight	Sales	Rating
0	Regular	FDX32	Fruits and Vegetables	2012	OUT049	Tier 1	Medium	Supermarket Type1	0.100014	15.10	145.4786	5.0
1	Low Fat	NCB42	Health and Hygiene	2022	OUT018	Tier 3	Medium	Supermarket Type2	0.008596	11.80	115.3492	5.0
2	Regular	FDR28	Frozen Foods	2016	OUT046	Tier 1	Small	Supermarket Type1	0.025896	13.85	165.0210	5.0
3	Regular	FDL50	Canned	2014	OUT013	Tier 3	High	Supermarket Type1	0.042278	12.15	126.5046	5.0
4	Low Fat	DRI25	Soft Drinks	2015	OUT045	Tier 2	Small	Supermarket Type1	0.033970	19.60	55.1614	5.0
...	...	...	...	...	...	...	...	...	...	...	...	...
8518	low fat	NCT53	Health and Hygiene	2018	OUT027	Tier 3	Medium	Supermarket Type3	0.000000	NaN	164.5526	4.0
8519	low fat	FDN09	Snack Foods	2018	OUT027	Tier 3	Medium	Supermarket Type3	0.034706	NaN	241.6828	4.0
8520	low fat	DRE13	Soft Drinks	2018	OUT027	Tier 3	Medium	Supermarket Type3	0.027571	NaN	86.6198	4.0
8521	reg	FDT50	Dairy	2018	OUT027	Tier 3	Medium	Supermarket Type3	0.107715	NaN	97.8752	4.0
8522	reg	FDM58	Snack Foods	2018	OUT027	Tier 3	Medium	Supermarket Type3	0.000000	NaN	112.2544	4.0

8523 rows × 12 columns

3.Data Cleaning

isnull().sum() : isnull().sum()in Pandas is used to check missing values in a dataset. isnull() marks each cell as True if it is empty (NaN) and **False** if it has data. Then .sum() adds up all the True values column-wise, giving the total number of missing values in each column. It is a quick way to find how much data is missing.

```
print(df.isnull().sum())
```

Python

Output:

```
Item Fat Content             0.000000
Item Identifier              0.000000
Item Type                    0.000000
Outlet Establishment Year    0.000000
Outlet Identifier            0.000000
Outlet Location Type         0.000000
Outlet Size                  0.000000
Outlet Type                  0.000000
Item Visibility              0.000000
Item Weight                  0.171653
Sales                        0.000000
Rating                       0.000000
dtype: float64
```

isnull().mean(): isnull().mean()in Pandas is used to find the fraction of missing values in each column. First, isnull() marks missing entries as True (1) and non-missing as False (0). Then, mean() calculates the average of these values, giving the proportion of nulls. For example, if a column has 100 rows and 25 are missing,the result will be 0.25, meaning 25% of the data is missing in that column.

```
print(df.isnull().mean())
```

✓ 0.0s

Output:

```
Item Fat Content                0
Item Identifier                 0
Item Type                       0
Outlet Establishment Year       0
Outlet Identifier               0
Outlet Location Type            0
Outlet Size                     0
Outlet Type                     0
Item Visibility                 0
Item Weight                  1463
Sales                           0
Rating                          0
dtype: int64
```

4.Options to clean

Drop rows/columns with too many missing values:

df.dropna(): df.dropna() is a pandas function that removes rows or columns from a DataFrame if they contain missing (NaN) values. By default, it drops any row that has at least one missing value. You can adjust it to drop columns instead, or only drop rows/columns when all values are missing. It helps clean the data by getting rid of incomplete information.

```python
df = df.dropna()
```

Python

5.Remove Duplicates

drop_duplicates():drop_duplicates()The function drop_duplicates() in Pandas is used to remove duplicate rows from a DataFrame. By default, it checks all columns, and if two or more rows have the same values, only the first occurrence is kept while the others are dropped. You can also specify a particular column to check for duplicates using the subset parameter.

This function helps in cleaning the dataset by ensuring that only unique records remain, making the data more accurate and reliable for analysis.

```python
print(df.drop_duplicates())
```

✓ 0.0s

Python

6. Fix Inconsistent Data

Replace(): This code replaces different spellings and short forms in the Item Fat Content column so that all values become consistent as either "Low Fat or "Regular."

```python
df['Item Fat Content'] = df['Item Fat Content'].replace({
    'LF':'Low Fat',
    'low fat':'Low Fat',
    'reg':'Regular'
})
df
```

Python

Output:

	Item Fat Content	Item Identifier	Item Type	Outlet Establishment Year	Outlet Identifier	Outlet Location Type	Outlet Size	Outlet Type	Item Visibility	Item Weight	Sales	Rating
0	Regular	FDX32	Fruits and Vegetables	2012	OUT049	Tier 1	Medium	Supermarket Type1	0.100014	15.10	145.4786	5.0
1	Low Fat	NCB42	Health and Hygiene	2022	OUT018	Tier 3	Medium	Supermarket Type2	0.008596	11.80	115.3492	5.0
2	Regular	FDR28	Frozen Foods	2016	OUT046	Tier 1	Small	Supermarket Type1	0.025896	13.85	165.0210	5.0
3	Regular	FDL50	Canned	2014	OUT013	Tier 3	High	Supermarket Type1	0.042278	12.15	126.5046	5.0
4	Low Fat	DRI25	Soft Drinks	2015	OUT045	Tier 2	Small	Supermarket Type1	0.033970	19.60	55.1614	5.0
...	...	...	...	...	...	...	...	...	...	...	...	...
8241	Regular	FDH26	Canned	2022	OUT018	Tier 3	Medium	Supermarket Type2	0.034841	19.25	141.1496	4.0
8242	Regular	FDG56	Fruits and Vegetables	2022	OUT018	Tier 3	Medium	Supermarket Type2	0.071744	13.30	59.7536	4.0
8243	Regular	FDM15	Meat	2022	OUT018	Tier 3	Medium	Supermarket Type2	0.057655	11.80	152.6366	4.0
8244	Regular	FDX57	Snack Foods	2022	OUT018	Tier 3	Medium	Supermarket Type2	0.047459	17.25	95.8068	4.0
8245	Regular	FDD10	Snack Foods	2022	OUT018	Tier 3	Medium	Supermarket Type2	0.046208	20.60	178.0344	4.0

7060 rows × 12 columns

7.Handling Missing Values

fillna():The **fillna()** function is used to replace missing values (NaN) in a column with a specific value, such as the column's mean, median, mode, or any fixed value.

```
df['Item Weight'] = df['Item Weight'].fillna(df['Item Weight'].mean())
df['Outlet Size'] = df['Outlet Size'].fillna(df['Outlet Size'].mode()[0])


print(df['Item Weight'])
```

✓ 0.0s Python

Output:

```
0       15.100000
1       11.800000
2       13.850000
3       12.150000
4       19.600000
          ...
8518    12.857645
8519    12.857645
8520    12.857645
8521    12.857645
8522    12.857645
Name: Item Weight, Length: 8523, dtype: float64
```

```
print(df['Outlet Size'])
```

✓ 0.0s Python

```
0       Medium
1       Medium
2        Small
3         High
4        Small
          ...
8518    Medium
8519    Medium
8520    Medium
8521    Medium
8522    Medium
Name: Outlet Size, Length: 8523, dtype: object
```

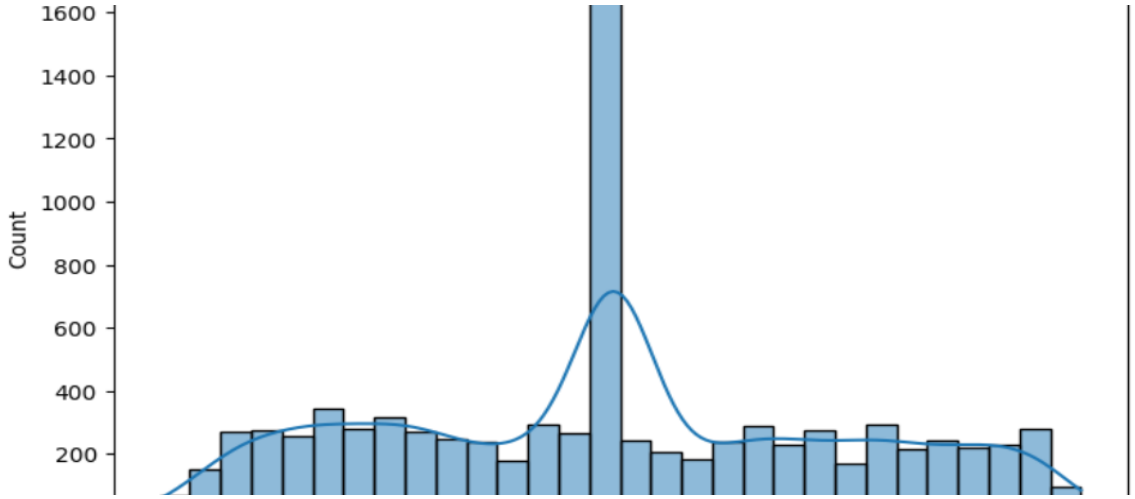
3.VISUALIZATION USING MATPLOTLIB & SEABORN

```python
import matplotlib.pyplot as plt
import seaborn as sns

plt.figure(figsize=(8,5))
sns.histplot(df['Item Weight'], bins=30, kde=True)
plt.title("Distribution of Item Weight")
plt.xlabel("Item Weight")
plt.ylabel("Count")
plt.show()
```

Python

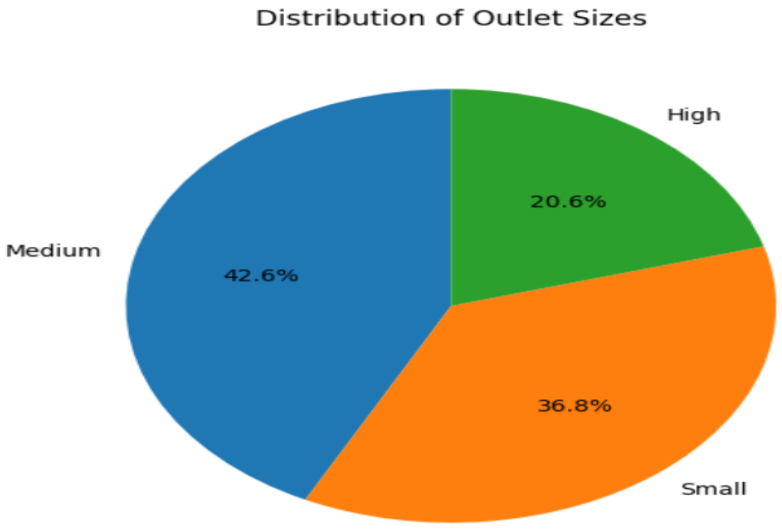
Output:



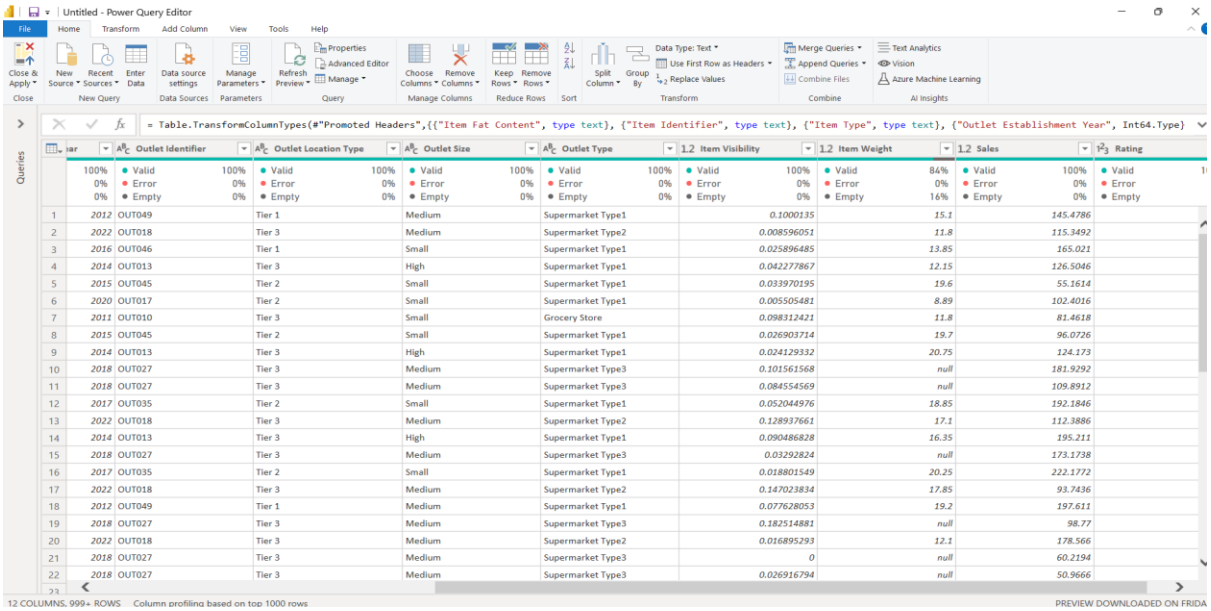

```python
import matplotlib.pyplot as plt
outlet_counts = df['Outlet Size'].value_counts()
plt.figure(figsize=(6,6))
plt.pie(outlet_counts, labels=outlet_counts.index, autopct='%1.1f%%', startangle=90)
plt.title("Distribution of Outlet Sizes")
plt.show()
```

Python

Output:



TRANSFORMATION IN POWER BI



4.DASHBOARD DESIGN AND VISUALIZATION

- An interactive dashboard was developed in Power BI to visually represent key insights.
- Visual elements such as bar charts, line charts, donut charts, and KPI cards were used.
- A filter panel was added to allow users to analyze sales trends across outlet size, type, and location.
- Sales performance was evaluated across different years, outlet types, and product categories.
- Customer purchase patterns were studied to understand preferences.
- Operational bottlenecks in outlet size and regional sales were identified.
- Final insights were documented to support recommendations for improving Blinkit’s sales and operational strategies.

Findings and Insights

The analysis of the Blinkit dataset through the Power BI dashboard revealed several important findings:

1. Sales Performance

- The total sales recorded were $1.20M, with an average sales value of $141.
- A total of 8,523 items were sold, and the average customer rating stood at 3.9, indicating moderate customer satisfaction.

2. Product Performance

- Regular fat items contributed significantly higher sales compared to low-fat items.
- Among product categories, fruits & vegetables, snack foods, and household items were the top contributors to revenue.

3. Outlet Analysis

- High-size outlets generated the highest sales compared to medium and small outlets.
- Tier 3 locations recorded maximum revenue, followed by Tier 2 and Tier 1.
- In terms of outlet types, Supermarket Type2 performed the best in sales, while grocery stores received the highest customer ratings.

4. Time-Based Trends

- Sales peaked around 2020, showing strong demand during that period, but declined in subsequent years.
- This highlights the impact of market conditions and consumer shifts on Blinkit's performance.

5. Operational Insights

- Variations in outlet size and regional demand suggest the need for optimized inventory distribution.
- Customer preferences for regular fat items and certain product categories indicate targeted marketing opportunities.
- The differences in outlet type performance suggest a scope to improve smaller outlets and grocery stores through better stock planning and promotions.

Recommendations

1. **Targeted Marketing**: Promote regular fat items and top-selling categories through personalized offers and discounts to boost sales further.

2. **Inventory Optimization**: Allocate more stock to high-demand categories such as fruits, vegetables, and snack foods, while reducing overstocking of low-performing items.

3. **Outlet Development**: Strengthen smaller and medium outlets by analyzing local demand and improving stock availability to match high-size outlets.

4. **Regional Strategy**: Focus on Tier 3 locations, which generate maximum revenue, and design region-specific campaigns to maintain growth.

5. **Customer Engagement**: Enhance customer satisfaction in outlets with lower ratings by improving service quality and product availability.

6. **Operational Improvement**: Streamline supply chain and delivery processes to reduce bottlenecks and ensure efficient order fulfillment.

7. **Data-Driven Forecasting:** Utilize historical sales data and seasonal trends to predict demand accurately, minimizing stockouts and overstock situations.

8. **Loyalty Programs:** Introduce or enhance customer loyalty initiatives to retain frequent buyers and encourage repeat purchases.

9. **Digital Presence Enhancement:** Expand online marketing, mobile app features, and social media engagement to attract tech-savvy customers and increase order frequency.

10. **Sustainability Initiatives:** Implement eco-friendly packaging and reduce food waste practices, which can improve brand image and appeal to environmentally conscious consumers.

11. **Training & Development:** Conduct regular training sessions for employees and outlet staff to improve customer service, product knowledge, and efficiency in handling operations.

OUTCOMES

1.Gained Practical Experience in the Data Analytics Lifecycle

- I worked through every stage of the data analytics process—data collection, cleaning, analysis, and visualization.
- This end-to-end exposure gave me a holistic understanding of how data moves through different phases before reaching a meaningful conclusion.

2.Strengthened Statistical Foundation for Decision-Making

- The application of statistical methods during the internship helped me transition from theoretical knowledge to real-world usage.
- It proved that sound statistics form the backbone of any reliable business decision.

3.Improved Python Coding Skills

- I gained the ability to write Python scripts for data analysis, including cleaning, visualization, and numerical calculations.
- These skills allow me to handle both structured (tables, CSV files) and unstructured data (text, logs, JSON files).

4.Ability to Contribute to Real-Time Business Projects

- Through exposure to real datasets and practical problem-solving, I gained confidence to contribute meaningfully to industry projects.
- I understood how theoretical learning must always be backed with practical application.

CONCLUSION

The internship at Alric Info Tech was an enriching and transformative experience that enabled me to bridge the gap between academic learning and industry practices. Over the course of thirty days, I was introduced to real-world applications of Data Analytics, which broadened my perspective beyond classroom theories and gave me a deeper appreciation of how data-driven decisions are made in professional settings.

One of the key highlights of this internship was the exposure I gained to industry-standard tools such as Excel, Power BI, and Python. Through practical assignments and hands-on projects, I learned how to clean, organize, and analyze raw datasets, extract meaningful patterns, and represent findings through effective visualizations. This not only strengthened my technical foundation but also helped me understand the importance of presenting data in a clear and actionable manner for decision-making.

Equally valuable was the opportunity to develop soft skills that are essential in any workplace. I learned how to collaborate effectively in teams, contribute ideas during discussions, and manage time while handling multiple tasks. The environment encouraged problem-solving and analytical thinking, which improved my ability to approach challenges logically and systematically. Additionally, regular interactions with mentors and peers helped me improve my communication skills, enabling me to explain complex findings in simple and structured ways.

Overall, this internship provided me with a balanced combination of technical proficiency and professional growth. It gave me the confidence to apply my academic knowledge in practical scenarios, enhanced my adaptability to industry standards, and nurtured skills that will be valuable throughout my career.

SUMMARY OF LEARNING

- Hands-on knowledge of Excel,Power BI, and Python libraries.
- Application of statistical techniques in practical scenarios.
- Development of interactive dashboards and reports for data storytelling.
- Understanding the complete data analytics lifecycle from raw data to actionable insights.

How it Helps in Career Goals?

This internship has played a vital role in shaping my career aspirations. The practical exposure gave me confidence to handle real-world datasets and apply analytical techniques effectively.

- It has prepared me to pursue a career in Data Analytics - the fastest-growing fields today.
- The skills gained—particularly in Python, BI tools, and statistics—are directly aligned with industry requirements.
- I now feel confident in my ability to work with large datasets, identify trends, and make data-driven decisions that can contribute to organizational growth.
- By working in a professional environment, I developed adaptability and teamwork skills, ensuring I can collaborate effectively and thrive in diverse workplace settings.
- The experience has enhanced my problem-solving and critical thinking abilities, which are essential for tackling complex business and technological challenges.

REFERENCES

Books Academic References

1. Albright, S. C., Winston, W. L., C Zappe, C. (2019). *Data Analysis and Decision Making*. Cengage Learning.
2. McKinney, W. (2018). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. O'Reilly Media.
3. Provost, F., C Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media.
4. Montgomery, D. C., C Runger, G. C. (2018). *Applied Statistics and Probability for Engineers*. Wiley.
5. Seema Acharya. (2019). *Data Analytics Using Python*. McGraw Hill Education.

WEBSITES &ONLINE RESOURCES

1.Microsoft Excel Documentation-https://support.microsoft.com/excel

2.Power BI Documentation – https://learn.microsoft.com/power-bi

3.Python Official Documentation – https://docs.python.org/3/

4.Pandas Library Documentation – https://pandas.pydata.org

5.NumPy Library Documentation – https://numpy.org

6.Matplotlib Official Documentation https://matplotlib.org/stable/index.html

7.Seaborn Documentation – https://seaborn.pydata.org

RESEARCH ARTICLES

1. Davenport, T. H., C Patil, D. J. (2012). *Data Scientist: The Sexiest Job of the 21st Century*. Harvard Business Review.
2. Manyika, J., et al. (2011). *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute.