High-Level Design (HLD)

ANALYZING AMAZON SALES DATA

PROHIT RATHOD

Document Version Control

Date Issued	Version	Description	Author
15/01/2023	1.0	Abstract, Introduction, General Description	Prohit Rathod
15/02/2023	1.1	Design Detail, KPI, Deployment	Prohit Rathod

Contents

ABSTRACT

Organizations under the E-commerce industry seek to attain core competence by creating and sustaining a unique process to collect personal information about customers and their purchasing trends. The report critically evaluates how a service-based organization - Amazon uses Management information systems as a vibrant tool in attaining competitive advantage through efficient management and acquisition of information. As in today's market without proper sales management, it's very hard to predict how the business is running and how it will be in future. Many companies with proper sales management have shown better growth as they already know which item they have to focus on, which product needs some improvement etc. Sales Management helps in maintaining its customer base for a longer time by providing them attractive offers, as they already have the information's like who are their top customers, whom they have to focus on etc. Sales Management also helps in minimizing the losses. Also, Competition is increasing day by day as many new companies are coming with better management systems and giving tough competition due to that it is now very important to have a proper Sales Management to run any business andto compete with these companies.

1 Introduction

1.1 Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions before coding and can be used as a reference manualfor how the modules interact at a high level.

The HLD will:

- Present all of the design aspects and define them in detail
- Describe the user interface being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and the architecture of the project
- List and describe the non-functional attributes like:

-Security
-Reliability
-Maintainability
-Portability
-Reusability
-Application compatibility
-Resource utilization
-Serviceability

1.2 Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

2 General Description

2.1 Product Perspective & Problem Statement

This Project aims to Analyze Amazon Sales from 2017 to 2019 to know more in-depth information about the items that are in high demand, items that are generating high profitswhich items should not be sold and how much stock we have to maintain for further Salesetc.

2.2 Tools used

Business Intelligence tools and libraries works such as NumPy, Pandas, Seaborn, Matplotlib, MS-Excel, MS-Power BI, Jupyter Notebook and Python Programming Languageare used to build the whole framework.

3 Design Details

3.1 Functional Architecture

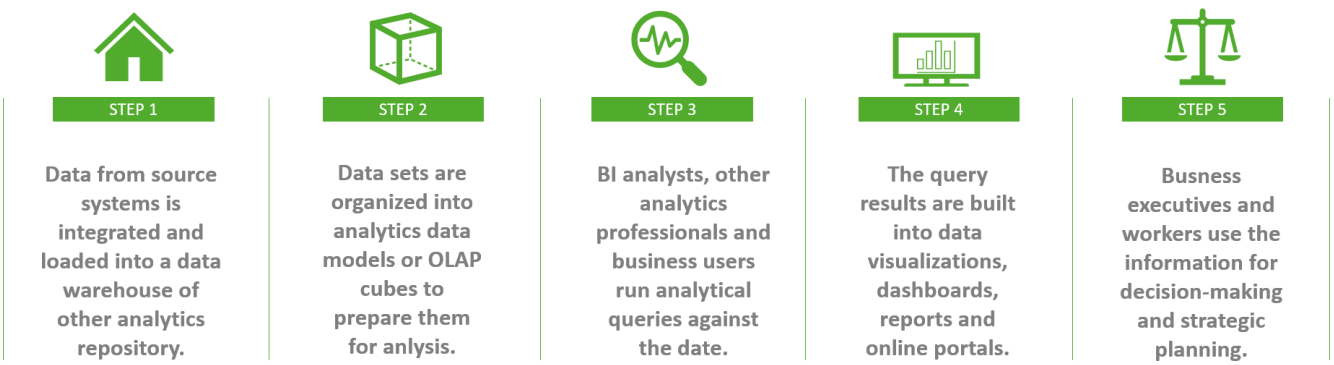
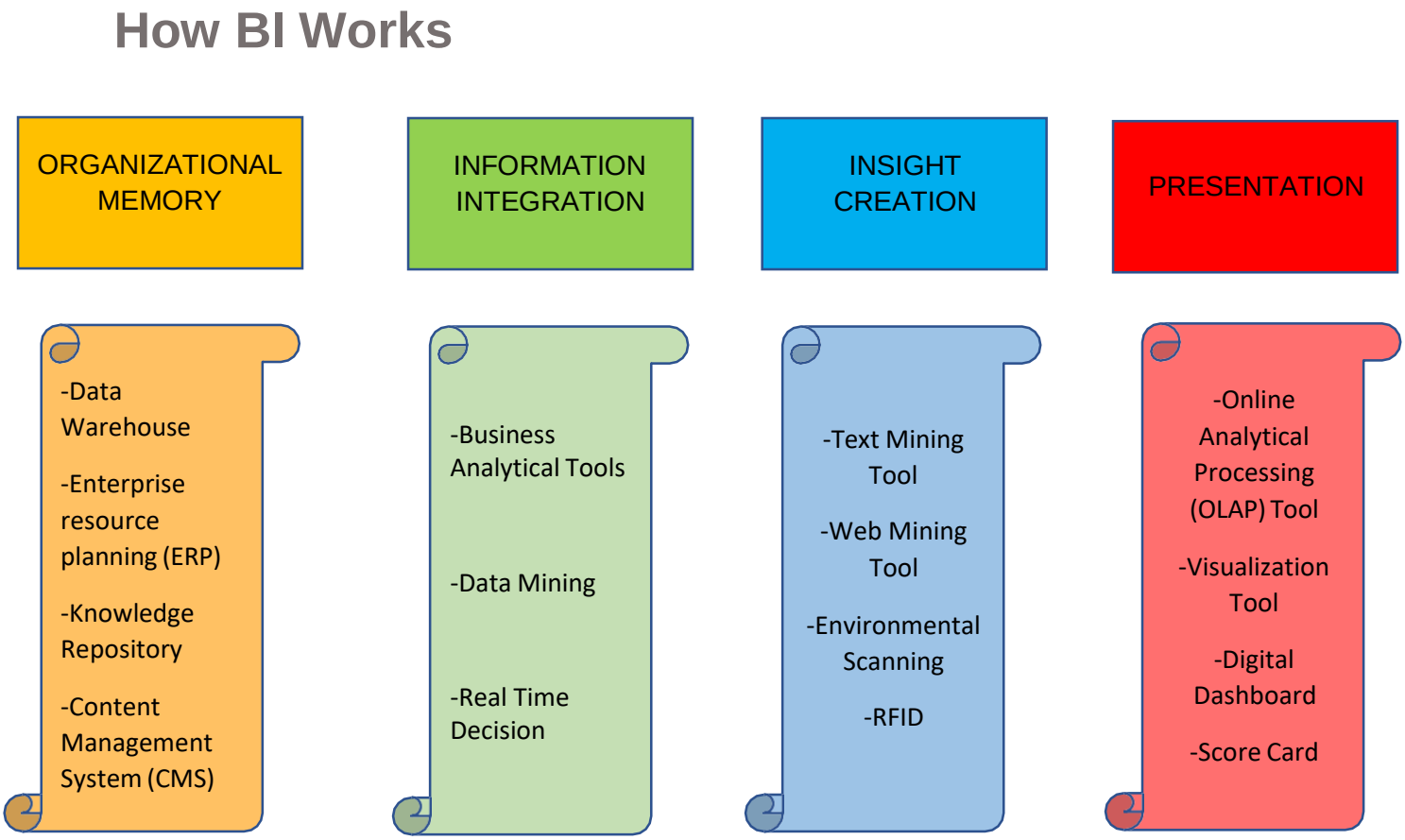


Figure 1: Functional Architecture of Business Intelligence



3.2 Optimization

i. Your data strategy drives performance

- Minimize the number of fields
- Minimize the number of records
- Optimize extracts to speed up future queries by materializing calculations,removing columns and the use of accelerated views

ii. Reduce the marks (data points) in your view

- Practice guided analytics. There's no need to fit everything you plan to show in a single view. Compile related views and connect them with action filters totravel from overview to highly-granular views at the speed of thought.
- Remove unneeded dimensions from the detail shelf.
- Explore. Try displaying your data in different types of views.

iii. Limit your filters by number and type

- Reduce the number of filters in use. Excessive filters on a view will create a more complex query, which takes longer to return results. Double-check yourfilters and remove any that aren't necessary.
- Use an include filter. Exclude filters load the entire domain of a dimensionwhile including filters do not. An include filter runs much faster than an exclude filter, especially for dimensions with many members.
- Use a continuous date filter. Continuous date filters (relative and range-of- date filters) can take advantage of the indexing properties in your databaseand are faster than discrete data filters.
- Use Boolean or numeric filters. Computers process integers and Booleans(t/f) much faster than strings.
- Use parameters and action filters. These reduce the query load (and workacross data sources).

iv. Optimize and materialize your calculations

- Perform calculations in the database
- Reduce the number of nested calculations.
- Reduce the granularity of LOD or table calculations in the view. The more granular the calculation, the longer it takes.

✓ LODs - Look at the number of unique dimension members in thecalculation.
✓ Table Calculations - the more marks in the view, the longer it will take tocalculate.

- Where possible, use MIN or MAX instead of AVG. AVG requires moreprocessing than MIN or MAX. Often rows will be duplicated and display thesame result with MIN, MAX, or AVG.
- Make groups with calculations. Like include filters, calculated groups load only named members of the domain, whereas Tableau's group function loadsthe entire domain.
- Use Booleans or numeric calculations instead of string calculations. Computers can process integers and Booleans (t/f) much faster than strings.

Boolean>Int>Float>Date>DateTime>String.

4 KPI

Dashboards will be implemented to display and indicate certain KPIs and relevantindicators for the disease.

As and when the system starts to capture the historical/periodic data for a user, thedashboards will be included to display charts over time with progress on various indicators or factors.

4.1 KPIs (Key Performance Indicators)

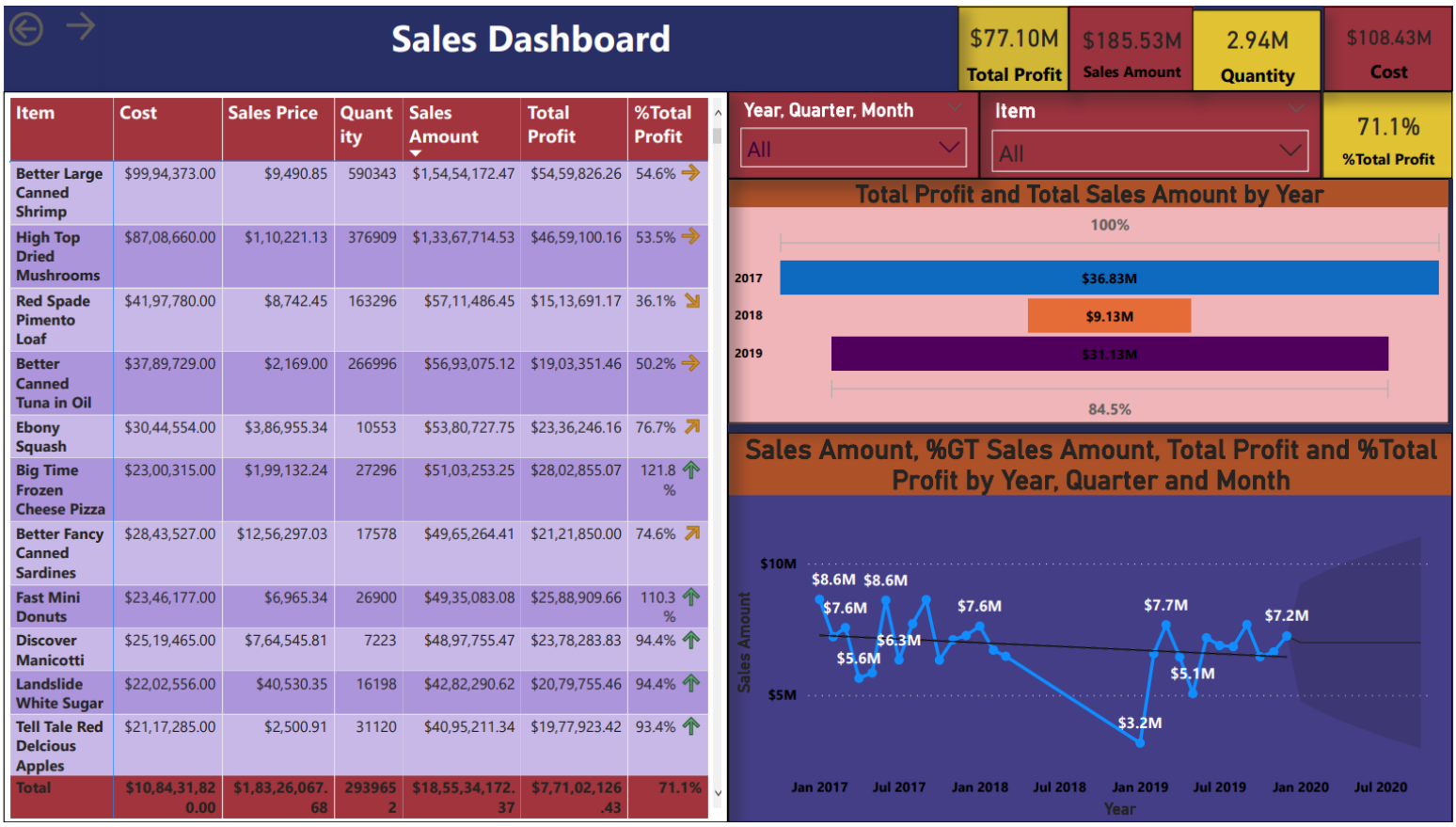
Key indicators displaying a summary of Sales Data and its relationships withdifferent metrics.

i. Yearly, Quarterly, Monthly Ups and Downs in Sales & Profits.
ii. Items That Generated Highest Sales, Profit etc.
iii. Top 5 Items that generated highest Sales and Top 5 Items by Quantity.
iv. Bottom 5 Items that generated Lowest Sales and Bottom 5 Items by Quantity.
v. Forecasting.

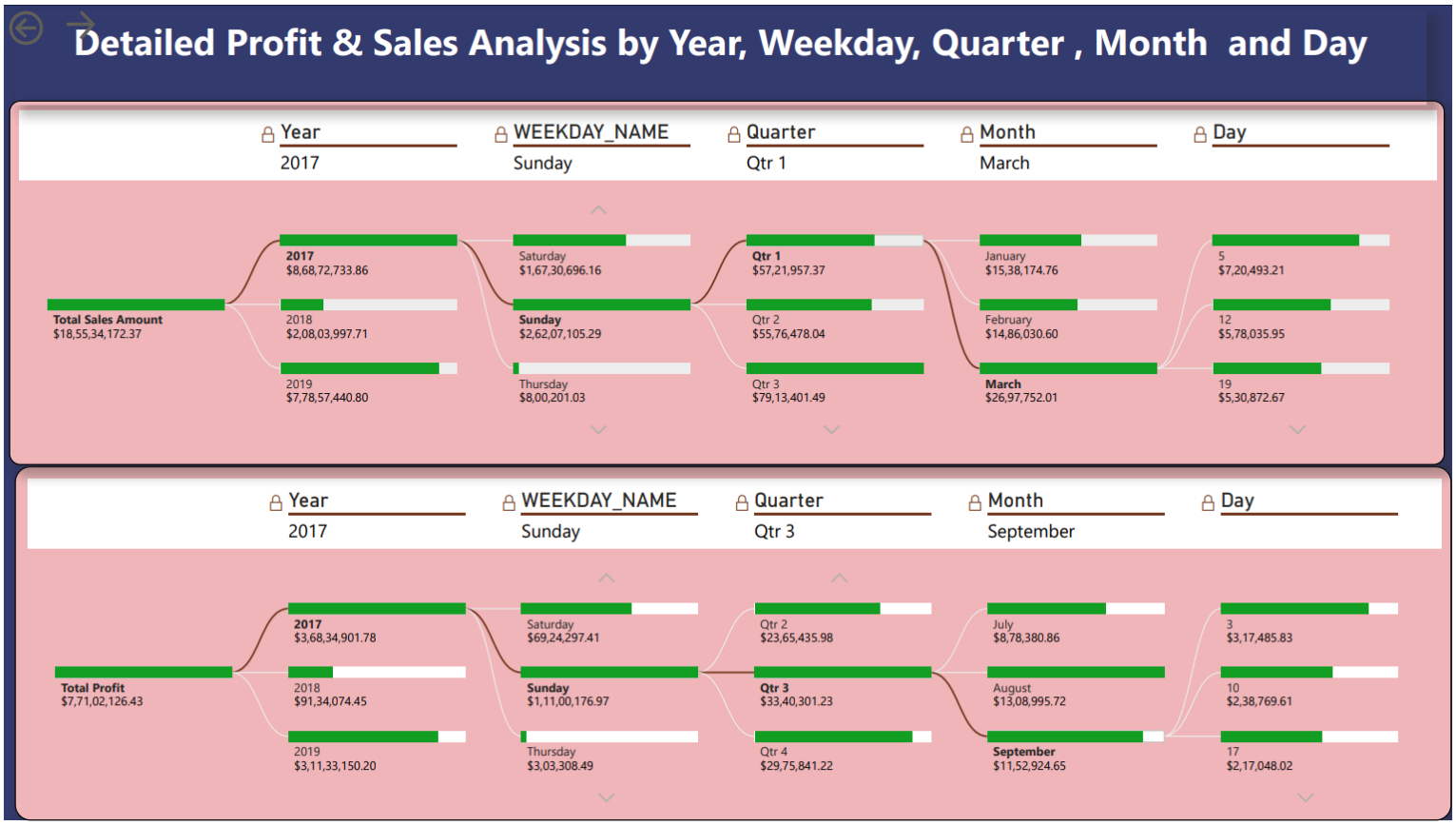
5 Deployment

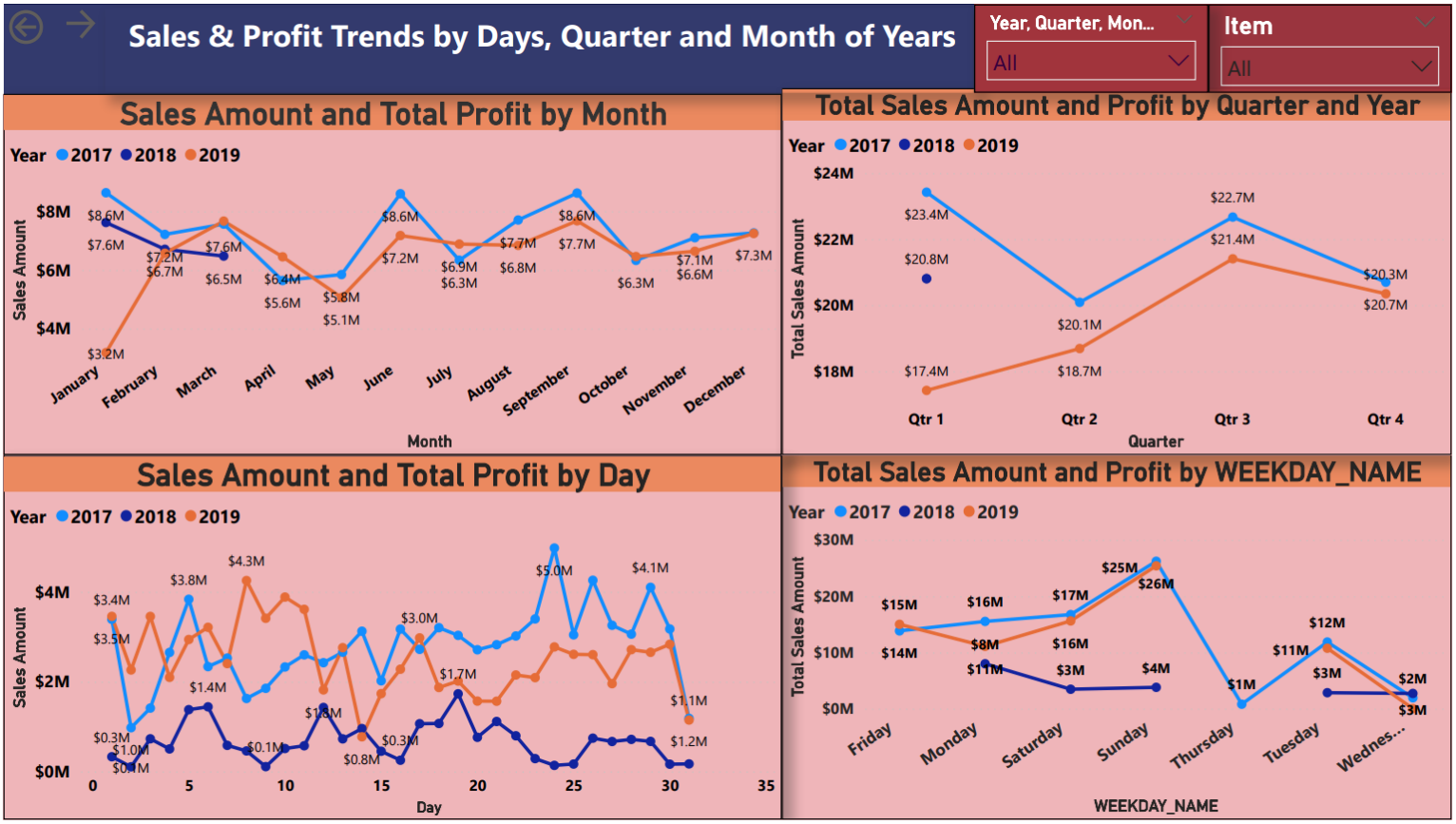
Prioritizing data and analytics couldn't come at a better time. Your company, no matter what size, is already collecting data and most likely analyzing just a portion of it to solvebusiness problems, gain competitive advantages, and drive enterprise transformation.
With the explosive growth of enterprise data, database technologies, and the high demand for analytical skills, today's most effective IT organizations have shifted their focus to enabling self-service by deploying and operating Power BI at scale, as well as organizing, orchestrating, and unifying disparate sources of data for business users andexperts alike to author and consume content.

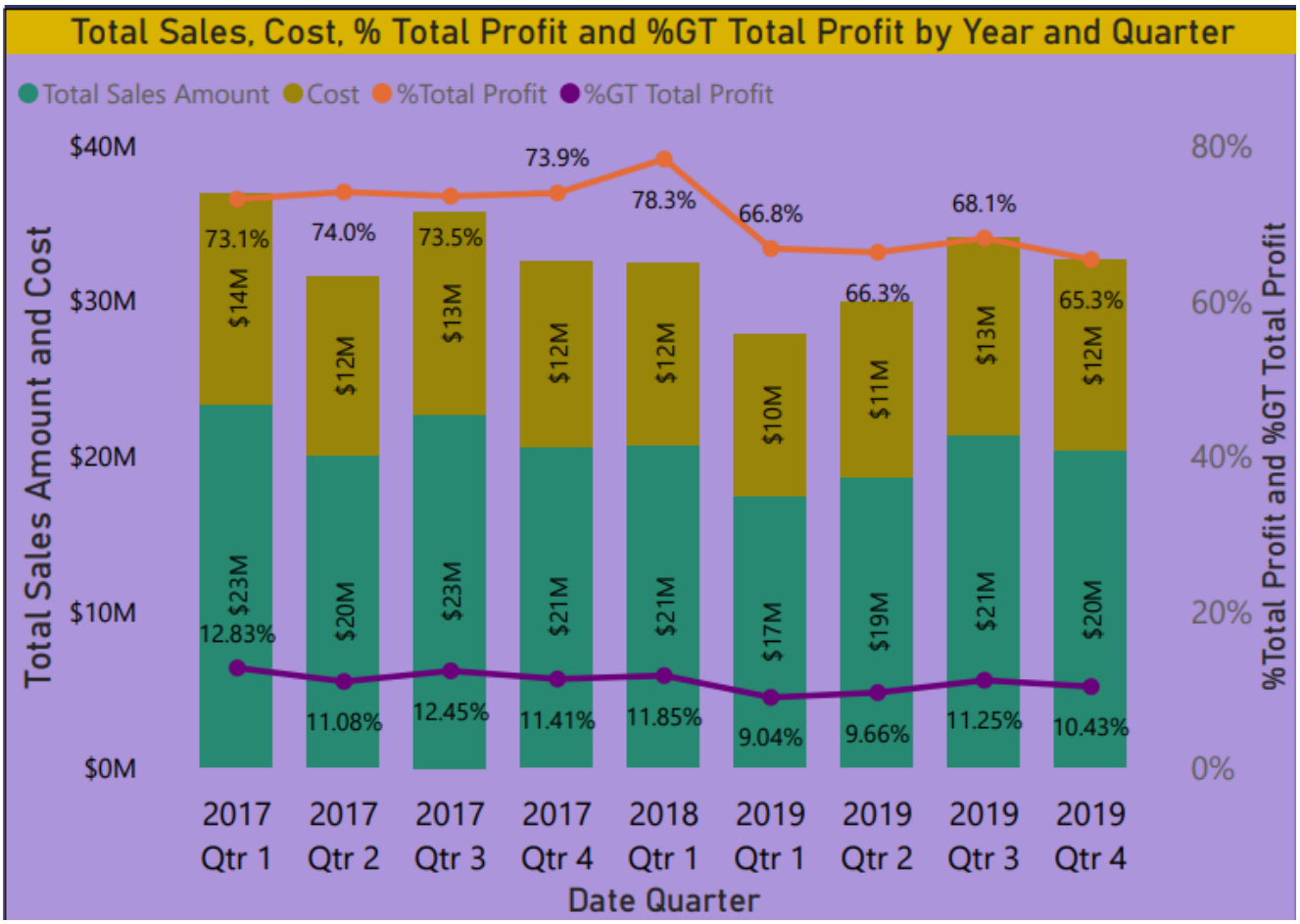
Power BI prioritizes choice in flexibility to fit, rather than dictate, your enterprise architecture. Power BI Desktop and Power BI Service leverage your existing technologyinvestments and integrate them into your IT infrastructure to provide a self-service, modern analytics platform for your users. With on-premises, cloud, and hosted options, there is a version of Power BI to match your requirements.

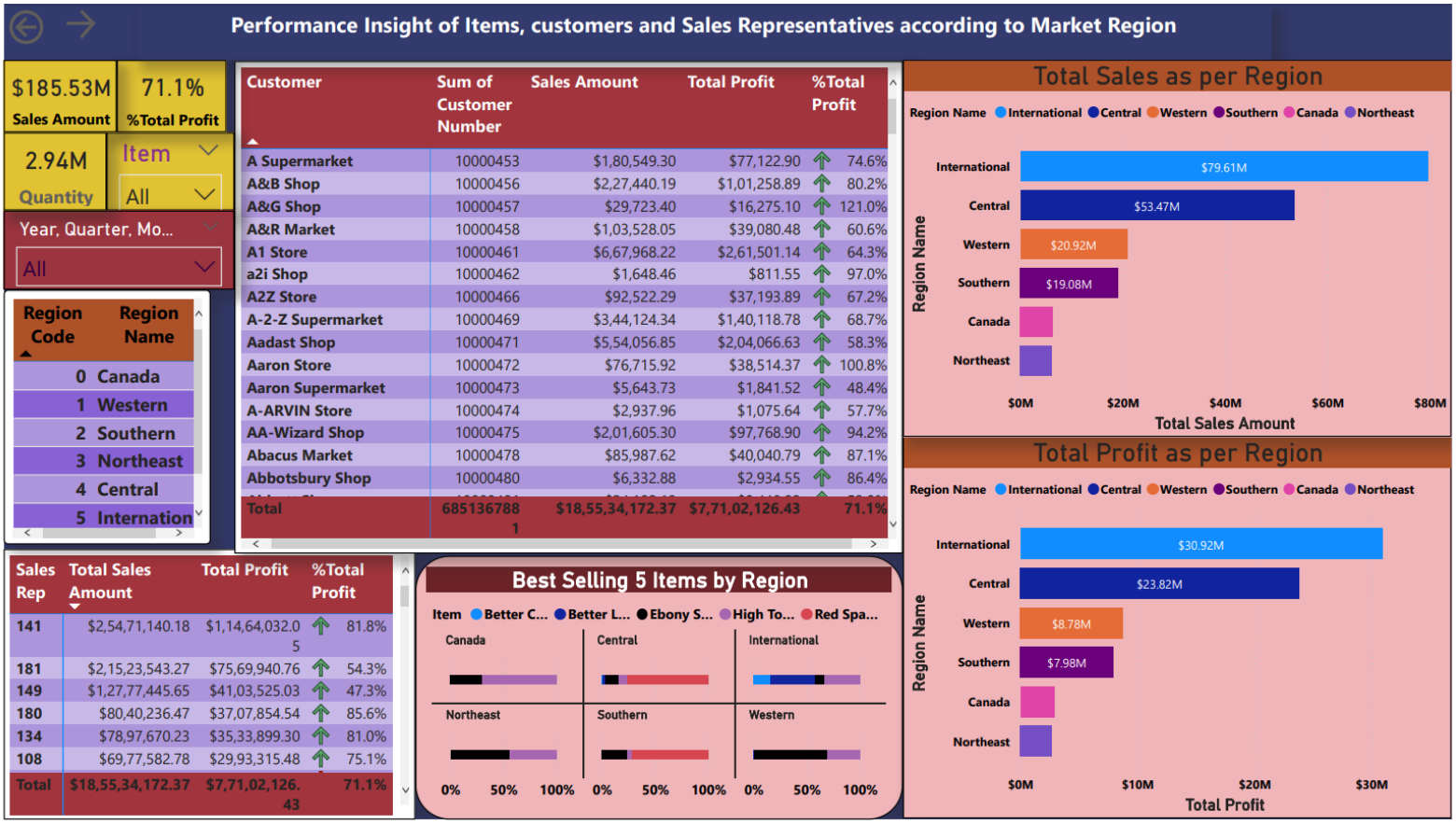


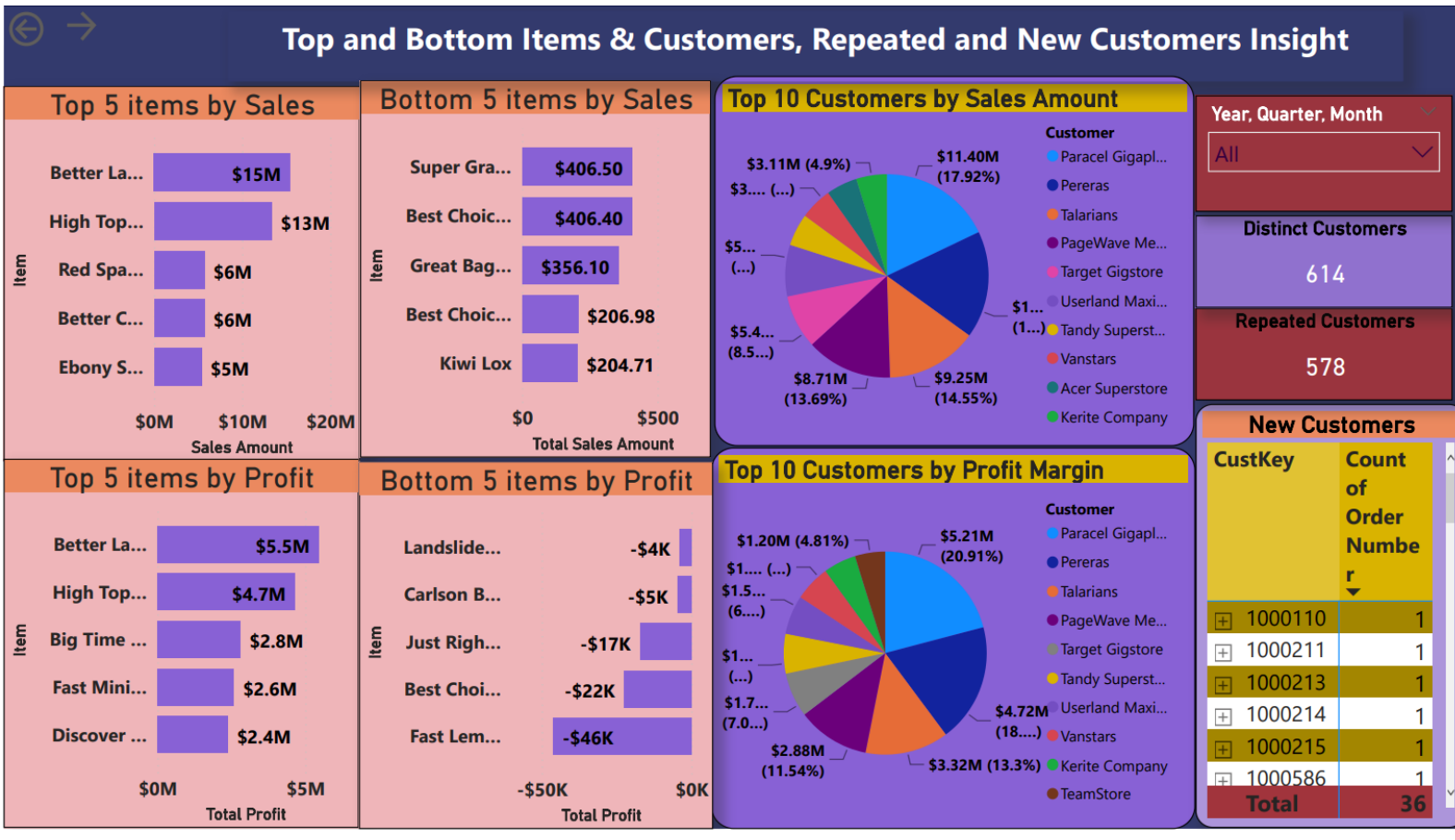
Item	Cost	Sales Price	Quantity	Sales Amount	Total Profit	%Total Profit
Better Large Canned Shrimp	$99,94,373.00	$9,490.85	590343	$1,54,54,172.47	$54,59,826.26	54.6%
High Top Dried Mushrooms	$87,08,660.00	$1,10,221.13	376909	$1,33,67,714.53	$46,59,100.16	53.5%
Red Spade Pimento Loaf	$41,97,780.00	$8,742.45	163296	$57,11,486.45	$15,13,691.17	36.1%
Better Canned Tuna in Oil	$37,89,729.00	$2,169.00	266996	$56,93,075.12	$19,03,351.46	50.2%
Ebony Squash	$30,44,554.00	$3,86,955.34	10553	$53,80,727.75	$23,36,246.16	76.7%
Big Time Frozen Cheese Pizza	$23,00,315.00	$1,99,132.24	27296	$51,03,253.25	$28,02,855.07	121.8 %
Better Fancy Canned Sardines	$28,43,527.00	$12,56,297.03	17578	$49,65,264.41	$21,21,850.00	74.6%
Fast Mini Donuts	$23,46,177.00	$6,965.34	26900	$49,35,083.08	$25,88,909.66	110.3 %
Discover Manicotti	$25,19,465.00	$7,64,545.81	7223	$48,97,755.47	$23,78,283.83	94.4%
Landslide White Sugar	$22,02,556.00	$40,530.35	16198	$42,82,290.62	$20,79,755.46	94.4%
Tell Tale Red Delcious Apples	$21,17,285.00	$2,500.91	31120	$40,95,211.34	$19,77,923.42	93.4%
Total	$10,84,31,820.00	$1,83,26,067.68	2939652	$18,55,34,172.37	$7,71,02,126.43	71.1%

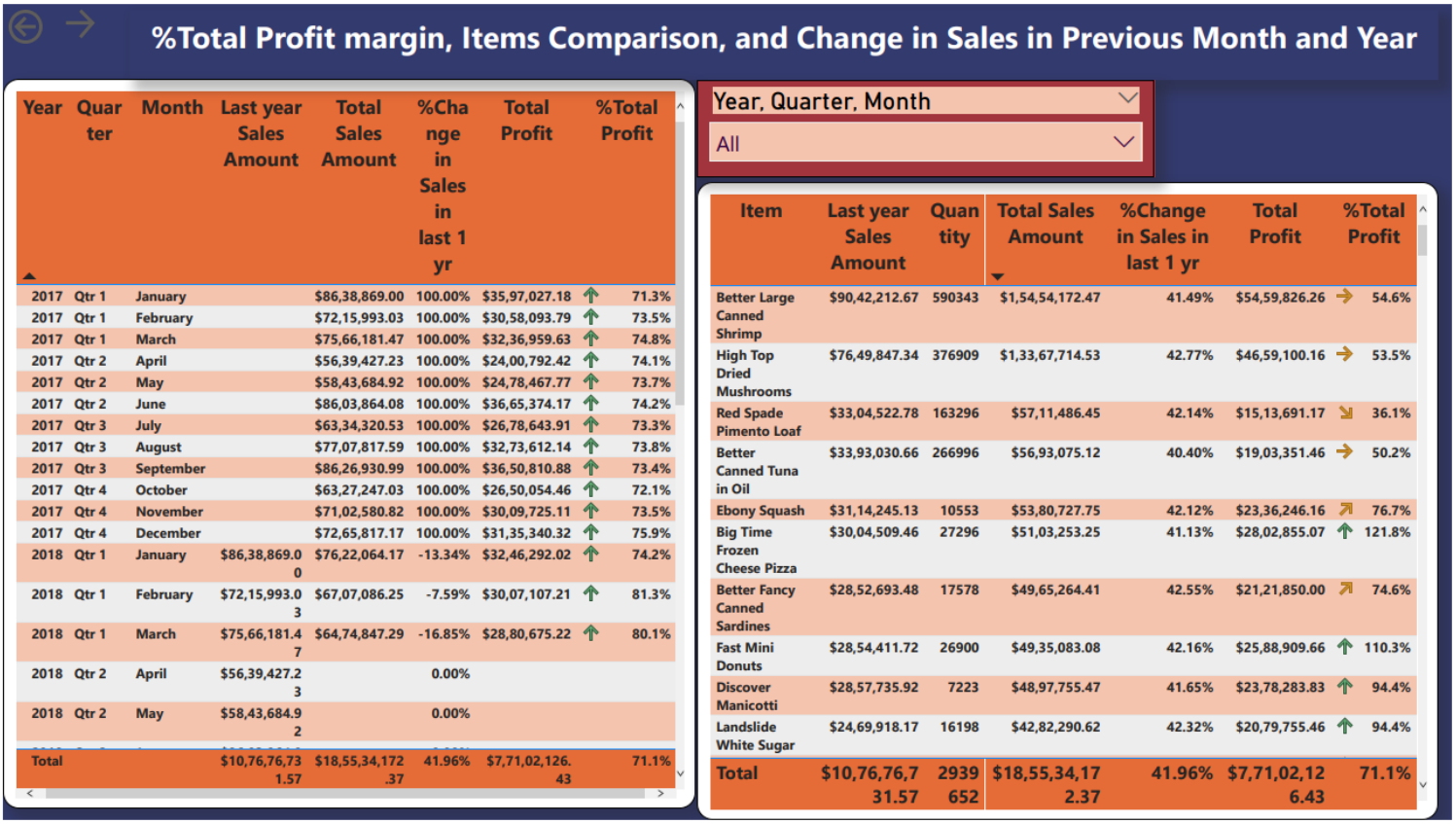



Sales & Profit Trends by Days, Quarter and Month of Years
Year, Quarter, Mon...
All
Item
All
Sales Amount and Total Profit by Month
Year 2017 2018 2019
Sales Amount
Month
January
February
March
April
May
June
July
August
September
October
November
December
Total Sales Amount and Profit by Quarter and Year
Year 2017 2018 2019
Total Sales Amount
Quarter
Qtr 1
Qtr 2
Qtr 3
Qtr 4
Sales Amount and Total Profit by Day
Year 2017 2018 2019
Sales Amount
Day
Total Sales Amount and Profit by WEEKDAY_NAME
Year 2017 2018 2019
Total Sales Amount
WEEKDAY_NAME
Friday
Monday
Saturday
Sunday
Thursday
Tuesday
Wednes...


Total Sales, Cost, % Total Profit and %GT Total Profit by Year and Quarter
Total Sales Amount
Cost
%Total Profit
%GT Total Profit
Total Sales Amount and Cost
%Total Profit and %GT Total Profit
Date Quarter
2017 Qtr 1
2017 Qtr 2
2017 Qtr 3
2017 Qtr 4
2018 Qtr 1
2019 Qtr 1
2019 Qtr 2
2019 Qtr 3
2019 Qtr 4


Performance Insight of Items, customers and Sales Representatives according to Market Region
$185.53M
Sales Amount
71.1%
%Total Profit
2.94M
Quantity
Item
All
Year, Quarter, Mo...
All
Region Code
Region Name
0 Canada
1 Western
2 Southern
3 Northeast
4 Central
5 Internation
Customer
Sum of Customer Number
Sales Amount
Total Profit
%Total Profit
A Supermarket 10000453 $1,80,549.30 $77,122.90 74.6%
A&B Shop 10000456 $2,27,440.19 $1,01,258.89 80.2%
A&G Shop 10000457 $29,723.40 $16,275.10 121.0%
A&R Market 10000458 $1,03,528.05 $39,080.48 60.6%
A1 Store 10000461 $6,67,968.22 $2,61,501.14 64.3%
a2i Shop 10000462 $1,648.46 $811.55 97.0%
A2Z Store 10000466 $92,522.29 $37,193.89 67.2%
A-2-Z Supermarket 10000469 $3,44,124.34 $1,40,118.78 68.7%
Aadast Shop 10000471 $5,54,056.85 $2,04,066.63 58.3%
Aaron Store 10000472 $76,715.92 $38,514.37 100.8%
Aaron Supermarket 10000473 $5,643.73 $1,841.52 48.4%
A-ARVIN Store 10000474 $2,937.96 $1,075.64 57.7%
AA-Wizard Shop 10000475 $2,01,605.30 $97,768.90 94.2%
Abacus Market 10000478 $85,987.62 $40,040.79 87.1%
Abbotsbury Shop 10000480 $6,332.88 $2,934.55 86.4%
Total 685136788 1 $18,55,34,172.37 $7,71,02,126.43 71.1%
Total Sales as per Region
Region Name International Central Western Southern Canada Northeast
International $79.61M
Central $53.47M
Western $20.92M
Southern $19.08M
Canada
Northeast
$0M $20M $40M $60M $80M
Total Sales Amount
Total Profit as per Region
Region Name International Central Western Southern Canada Northeast
International $30.92M
Central $23.82M
Western $8.78M
Southern $7.98M
Canada
Northeast
$0M $10M $20M $30M
Total Profit
Sales Rep
Total Sales Amount
Total Profit
%Total Profit
141 $2,54,71,140.18 $1,14,64,032.05 81.8%
181 $2,15,23,543.27 $75,69,940.76 54.3%
149 $1,27,77,445.65 $41,03,525.03 47.3%
180 $80,40,236.47 $37,07,854.54 85.6%
134 $78,97,670.23 $35,33,899.30 81.0%
108 $69,77,582.78 $29,93,315.48 75.1%
Total $18,55,34,172.37 $7,71,02,126.43 71.1%
Best Selling 5 Items by Region
Item Better C... Better L... Ebony S... High To... Red Spa...
Canada
Central
International
Northeast
Southern
Western
0% 50% 100%


Top and Bottom Items & Customers, Repeated and New Customers Insight
Top 5 items by Sales
Better La... $15M
High Top... $13M
Red Spa... $6M
Better C... $6M
Ebony S... $5M
Item
$0M $10M $20M
Sales Amount
Bottom 5 items by Sales
Super Gra... $406.50
Best Choic... $406.40
Great Bag... $356.10
Best Choic... $206.98
Kiwi Lox $204.71
Item
$0 $500
Total Sales Amount
Top 10 Customers by Sales Amount
$11.40M (17.92%)
$9.25M (14.55%)
$8.71M (13.69%)
$5.4... (8.5...)
$5... (...)
$3.... (...)
$3.11M (4.9%)
$1... (1...)
Customer
Paracel Gigapl...
Pereras
Talarians
PageWave Me...
Target Gigstore
Userland Maxi...
Tandy Superst...
Vanstars
Acer Superstore
Kerite Company
Year, Quarter, Month
All
Distinct Customers
614
Repeated Customers
578
New Customers
CustKey
Count of Order Number
1000110 1
1000211 1
1000213 1
1000214 1
1000215 1
1000586 1
Total 36
Top 5 items by Profit
Better La... $5.5M
High Top... $4.7M
Big Time ... $2.8M
Fast Mini... $2.6M
Discover ... $2.4M
Item
$0M $5M
Total Profit
Bottom 5 items by Profit
Landslide... -$4K
Carlson B... -$5K
Just Righ... -$17K
Best Choi... -$22K
Fast Lem... -$46K
Item
-$50K $0K
Total Profit
Top 10 Customers by Profit Margin
$5.21M (20.91%)
$4.72M (18....)
$3.32M (13.3%)
$2.88M (11.54%)
$1.7... (7.0...)
$1... (...)
$1.5... (6....)
$1.... (...)
$1.20M (4.81%)
Customer
Paracel Gigapl...
Pereras
Talarians
PageWave Me...
Target Gigstore
Tandy Superst...
Userland Maxi...
Vanstars
Kerite Company
TeamStore



Year	Quarter	Month	Last year Sales Amount	Total Sales Amount	%Change in Sales in last 1 yr	Total Profit	%Total Profit
2017	Qtr 1	January		$86,38,869.00	100.00%	$35,97,027.18	71.3%
2017	Qtr 1	February		$72,15,993.03	100.00%	$30,58,093.79	73.5%
2017	Qtr 1	March		$75,66,181.47	100.00%	$32,36,959.63	74.8%
2017	Qtr 2	April		$56,39,427.23	100.00%	$24,00,792.42	74.1%
2017	Qtr 2	May		$58,43,684.92	100.00%	$24,78,467.77	73.7%
2017	Qtr 2	June		$86,03,864.08	100.00%	$36,65,374.17	74.2%
2017	Qtr 3	July		$63,34,320.53	100.00%	$26,78,643.91	73.3%
2017	Qtr 3	August		$77,07,817.59	100.00%	$32,73,612.14	73.8%
2017	Qtr 3	September		$86,26,930.99	100.00%	$36,50,810.88	73.4%
2017	Qtr 4	October		$63,27,247.03	100.00%	$26,50,054.46	72.1%
2017	Qtr 4	November		$71,02,580.82	100.00%	$30,09,725.11	73.5%
2017	Qtr 4	December		$72,65,817.17	100.00%	$31,35,340.32	75.9%
2018	Qtr 1	January	$86,38,869.00	$76,22,064.17	-13.34%	$32,46,292.02	74.2%
2018	Qtr 1	February	$72,15,993.03	$67,07,086.25	-7.59%	$30,07,107.21	81.3%
2018	Qtr 1	March	$75,66,181.47	$64,74,847.29	-16.85%	$28,80,675.22	80.1%
2018	Qtr 2	April	$56,39,427.23		0.00%		
2018	Qtr 2	May	$58,43,684.92		0.00%		
Total			$10,76,76,731.57	$18,55,34,172.37	41.96%	$7,71,02,126.43	71.1%

Item	Last year Sales Amount	Quantity	Total Sales Amount	%Change in Sales in last 1 yr	Total Profit	%Total Profit
Better Large Canned Shrimp	$90,42,212.67	590343	$1,54,54,172.47	41.49%	$54,59,826.26	54.6%
High Top Dried Mushrooms	$76,49,847.34	376909	$1,33,67,714.53	42.77%	$46,59,100.16	53.5%
Red Spade Pimento Loaf	$33,04,522.78	163296	$57,11,486.45	42.14%	$15,13,691.17	36.1%
Better Canned Tuna in Oil	$33,93,030.66	266996	$56,93,075.12	40.40%	$19,03,351.46	50.2%
Ebony Squash	$31,14,245.13	10553	$53,80,727.75	42.12%	$23,36,246.16	76.7%
Big Time Frozen Cheese Pizza	$30,04,509.46	27296	$51,03,253.25	41.13%	$28,02,855.07	121.8%
Better Fancy Canned Sardines	$28,52,693.48	17578	$49,65,264.41	42.55%	$21,21,850.00	74.6%
Fast Mini Donuts	$28,54,411.72	26900	$49,35,083.08	42.16%	$25,88,909.66	110.3%
Discover Manicotti	$28,57,735.92	7223	$48,97,755.47	41.65%	$23,78,283.83	94.4%
Landslide White Sugar	$24,69,918.17	16198	$42,82,290.62	42.32%	$20,79,755.46	94.4%
Total	$10,76,76,731.57	2939652	$18,55,34,172.37	41.96%	$7,71,02,126.43	71.1%