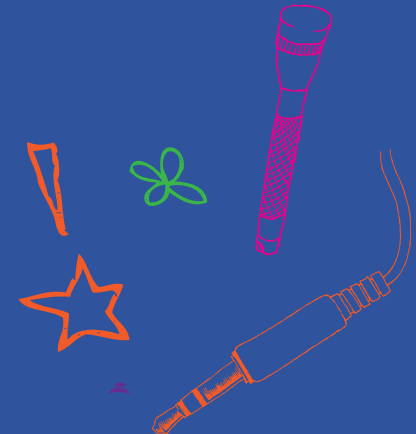


Analytics Transition Year Module 3



High performance. Delivered.



Overview of Project



	Module 1	Module 2	Module 3	Module 4
Overview	Introduction to the challenge	Using Analytics to understand the problem	Interpretation of Analysis	New Company Strategy and Business Report
Learning Outcomes	<ol style="list-style-type: none">1. Background of the company2. The Business Plan	<ol style="list-style-type: none">1. Understanding Data2. Applying Analytics	<ol style="list-style-type: none">1. Statistical Techniques2. Market Segmentation	<ol style="list-style-type: none">1. Forming a Strategy2. Completing a business report



Module 2 findings: Summary

In this section we will try to pull together all the insights we have found thus far in this case study.

We have to ask ourselves the following questions:

- ✓ What months are best for sales?
- ✓ What time periods are we selling most in?
- ✓ Which locations are we selling most in?
- ✓ Which vendors are the most profitable for the company currently?
- ✓ Which products are we selling most of?
- ✓ Which product sizes are doing best in the market place?

Data insights from the graphs produced

Vending machines and Mace

Question	Answer
Best Months	March September December
Best Times	12:00 – 13:00 19:00 – 22:00
Best Location	Clare
Best Vendor	Vending Machine
Best Product	Choco Milk
Best Product Size	Small

Corner Shops

Question	Answer
Best Month	March April December
Best Times	N/A
Best Location	Clare
Best Vendor	Corner Shop
Best Product	Choco Milk
Best Product Size	Small

Applying Statistical Techniques: Centrality

MEAN

Use the mean to describe the middle of a set of data that does not have an outlier.

Example:

If we were looking at sales per month the mean would be able to tell us the average number of sales per month for the selected data

MEDIAN

Use the median to describe the middle of a set of data that does have an outlier.

Example:

If we were looking at sales per month the median would help us to understand which shops are in the bottom half for sales and which shops are in the top half for sales

MODE

Use the mode when the data is non-numeric or when asked to choose the most popular item.

Example:

If we were looking at sales per month the mode would tell us the number of sales which is most common across the months and we would be able to see for how many months this number of sales happened

Applying Statistical Techniques: Centrality

Let's now calculate the **mean** for the number of bars sold in one transaction for Vending machine and Mace

$$\text{Mean} = \frac{\text{Total Number of Bars Sold}}{\text{Total Number of Transactions}}$$

$$= \frac{3+1+2+1+3+2+3+2+1+2+1+2+2+3+3+1+1+3+3+2+3+3+1+1+3}{25}$$

$$= \frac{52}{25} = 2.08 \text{ bars per transaction}$$

Applying Statistical Techniques: Centrality

Let's now calculate the **median** for the number of bars sold in one transaction for vending machine and Mace

Median is the middle number of the ordered data

Order Data:

1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,3,3,3,3,3,3,3,3,3,3

We have 25 pieces of data so the median will be at position 13 in our list above

Median = 2

Applying Statistical Techniques: Centrality

Let's now calculate the **mode** for the number of bars sold in one transaction

Mode is the number that appears most in the data

1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,3,3,3,3,3,3,3,3,3,3

Number of 1's = 8

Number of 2's = 7

Number of 3's = 10

Mode= 3

Applying Statistical Techniques: Full Dataset

We will now use Excel to see the values of the mean, median and mode for the entire vending machine and Mace data set.

Mean = 1.99

Median = 2

Mode = 1

Should we use the values we calculated ourselves or the ones above that we calculated using excel? Why?

Let's do some more calculations on other aspects of the data.

Applying Statistical Techniques: Centrality

Let's now calculate the **mean** for the number of boxes sold in one transaction for Corner Shops

$$\text{Mean} = \frac{\text{Total Number of Boxes Sold}}{\text{Total Number of Transactions}}$$

$$= \frac{1+3+2+1+1+2+2+1+1+1+1+3+2+3+2+1+1+3+1+3+3+3+1+3}{25}$$

$$= \frac{48}{25} = 1.92 \text{ boxes per transaction}$$

Applying Statistical Techniques: Centrality

Let's now calculate the **median** for the number of bars sold in one transaction for Corner Shops

Median is the middle number of the ordered data

Order Data:

1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,3,3,3,3

We have 25 pieces of data so the median will be at position 13 in our list above

Median = 2

Applying Statistical Techniques: Centrality

Let's now calculate the **mode** for the number of bars sold in one transaction

Mode is the number that appears most in the data

1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,3,3,3,3

Number of 1's = 11

Number of 2's = 5

Number of 3's = 9

Mode= 1

Applying Statistical Techniques: Full Dataset

We will now use Excel to see the values of the mean, median and mode for the entire vending machine and Mace data set.

Mean = 1.99

Median = 2

Mode = 1

Should we use the values we calculated ourselves or the ones above that we calculated using excel? Why?

Results of Statistical Techniques

- We also found the mean, median and mode for the number of bars sold by Vending machines and Mace shops in week 2
- We carried out the same process for the corner shop data

Measure	Vending Machines & Mace	Corner Shops
Mean	1.99	1.99
Median	2	2
Mode	1	1

Advance Analytics - Scatter plots

A scatter plot was used to determine the relationship between the corner shops monthly revenue and the number of boxes they purchased over the year

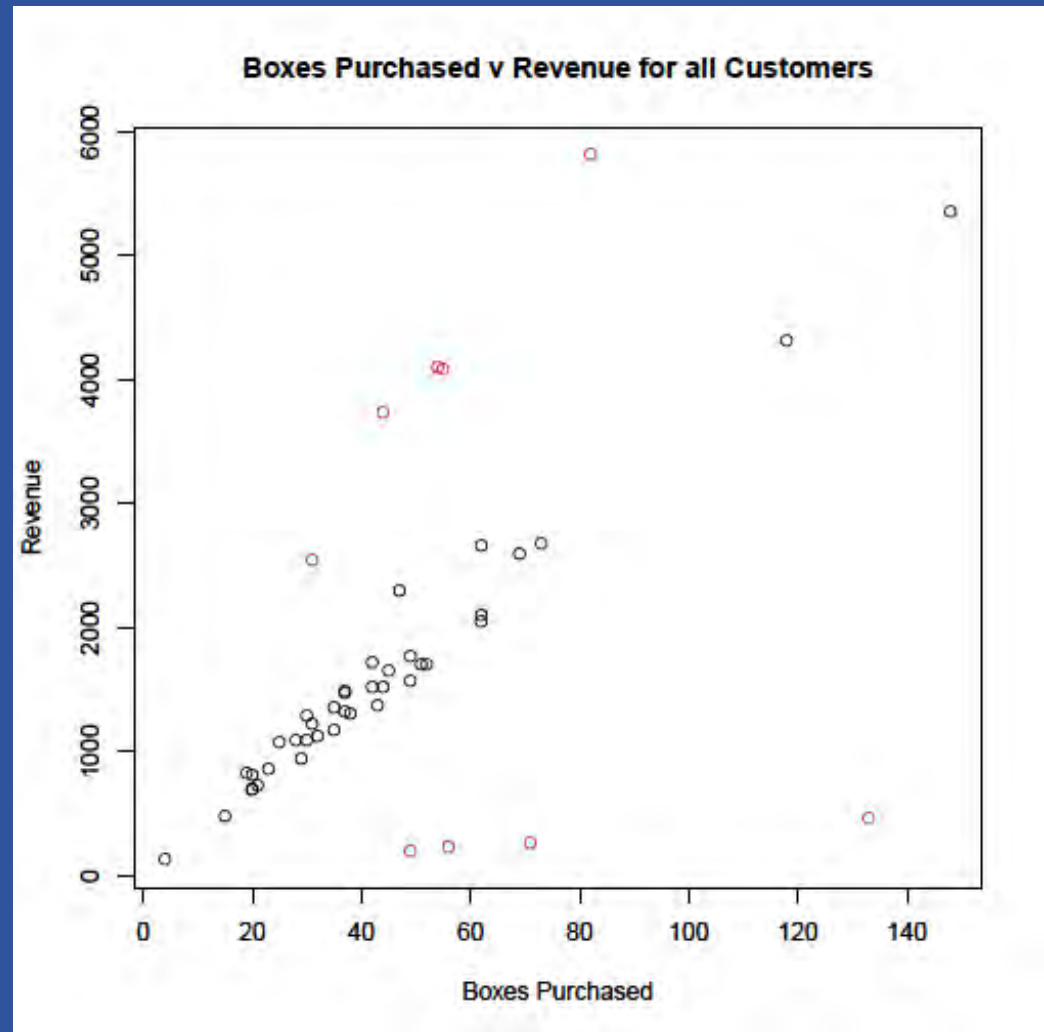
Questions

Is the relationship between revenue and purchases what you expected?

Can you explain the high outliers, points in red?

Can you explain the low outliers?

Which level of business, in terms of revenue, should Seacláid Ltd try and increase?



Short Analytics Puzzle

As we're about to start onto **market segmentation**, here's a short segmenting puzzle.

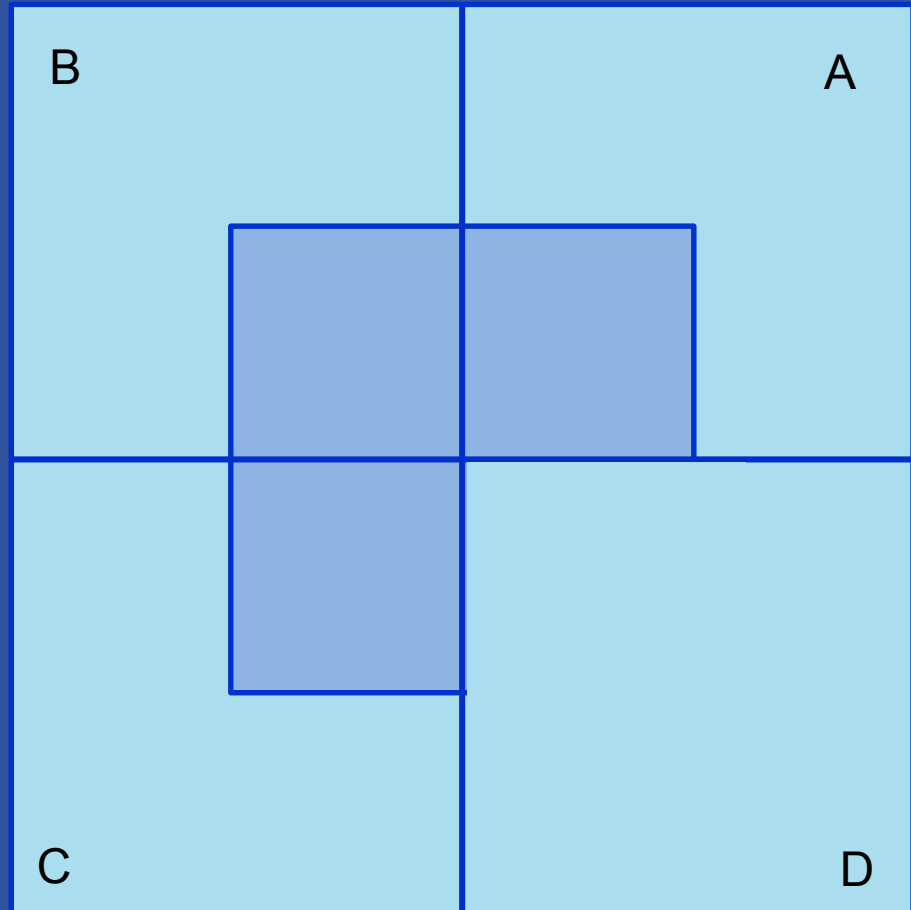
The light Shaded region is the area you'll need to segment.

For **Square A** – Split into **2** equal shapes.

For **Square B** – Split into **3** equal shapes.

For **Square C** – Split into **4** equal shapes, they're getting trickier

For **Square D** – Split into **7** equal shapes!!



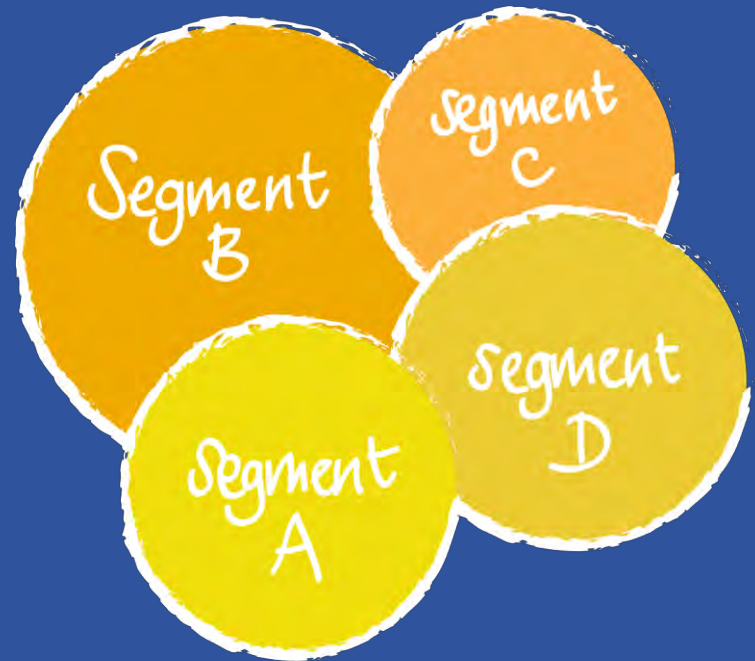
Objective 2: Market Segmentation

Market Segmentation

What is Market Segmentation?

Market segmentation is the process of dividing the customers into areas based on different characteristics.

It is about identifying the specific needs and wants of customer groups and then using those insights to provide products and services which meet customer needs.



Market Segmentation: The "Who"



In module 2 we discussed the *when, where, and what* for sales

We will now look at the *who* element and see who are buying Seaclaid's products

The data for this section only exists for the Mace shops because neither the vending machines or the corner shops can keep a detailed record of consumer category

We will first calculate the percentages of OAP's, Student's and Standard people who purchase the products from Mace shops, we will later use these percentages to segment the data into three unique subsets

Market Segmentation: Sample Data

Sample Data

Sold by	Consumer Category
Mace	Standard
Mace	Standard
Mace	OAP
Mace	Student
Mace	Standard
Mace	OAP
Mace	Student
Mace	OAP
Mace	Student
Mace	Standard
Mace	Student
Mace	Standard
Mace	Student
Mace	OAP
Mace	Standard
Mace	Standard
Mace	Student
Mace	Student
Mace	Student
Mace	OAP
Mace	OAP
Mace	Student
Mace	OAP
Mace	Student
Mace	Standard

Formula

$$\% \text{ of type in Sample} = \frac{\text{Total of Type}}{\text{Total Observations}}$$

Calculations

Number of OAPs: 7

Number of Students: 10

Number of Standard: 8

Total Number of Observations: 25

$$\% \text{ of OAPs} = \frac{7}{25} = 28\%$$

$$\% \text{ of Students} = \frac{10}{25} = 40\%$$

$$\% \text{ of Standard} = \frac{8}{25} = 32\%$$

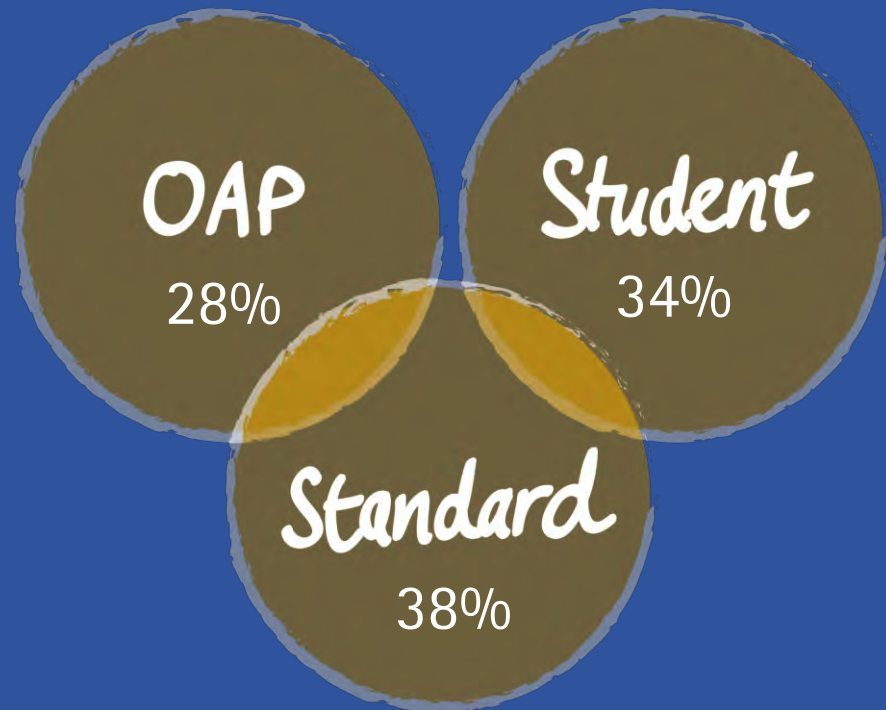
Market Segmentation: Full Dataset

We will know calculate the % of each type of customer in the dataset using Excel

% of OAP's = 28%

% of Students = 34%

% of Standard = 38%



Predictive Analytics

From the % that we just found relating to *who* is buying our products we can now apply some basic predictive analytics to see who is going to buy our products



From the probabilities that we found, which customer type is most likely to buy a Seaclaid Ltd. product?

If we knew that 2 people would buy Seacláid Ltd products on a given day, how many permutations of customer type are possible if replacement is allowed?

Now calculate the probability that the two customers that arrive are from the same category

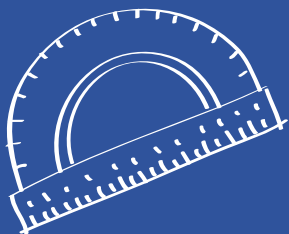


Predictive Analytics

Somebody from the Standard Customer category is most likely to buy a Seaclaid Ltd. product since the % of this category is greatest in the population

Permutations with Replacement

Student, Standard	Standard, Student	OAP, Student
Student, OAP	Standard, OAP	OAP, Standard
Student, Student	Standard, Standard	OAP, OAP

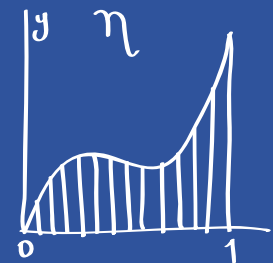


$$P(\textit{Student, Student}) = 0.34 * 0.34 = 0.1156$$

$$P(\textit{Standard, Standard}) = 0.38 * 0.38 = 0.1444$$

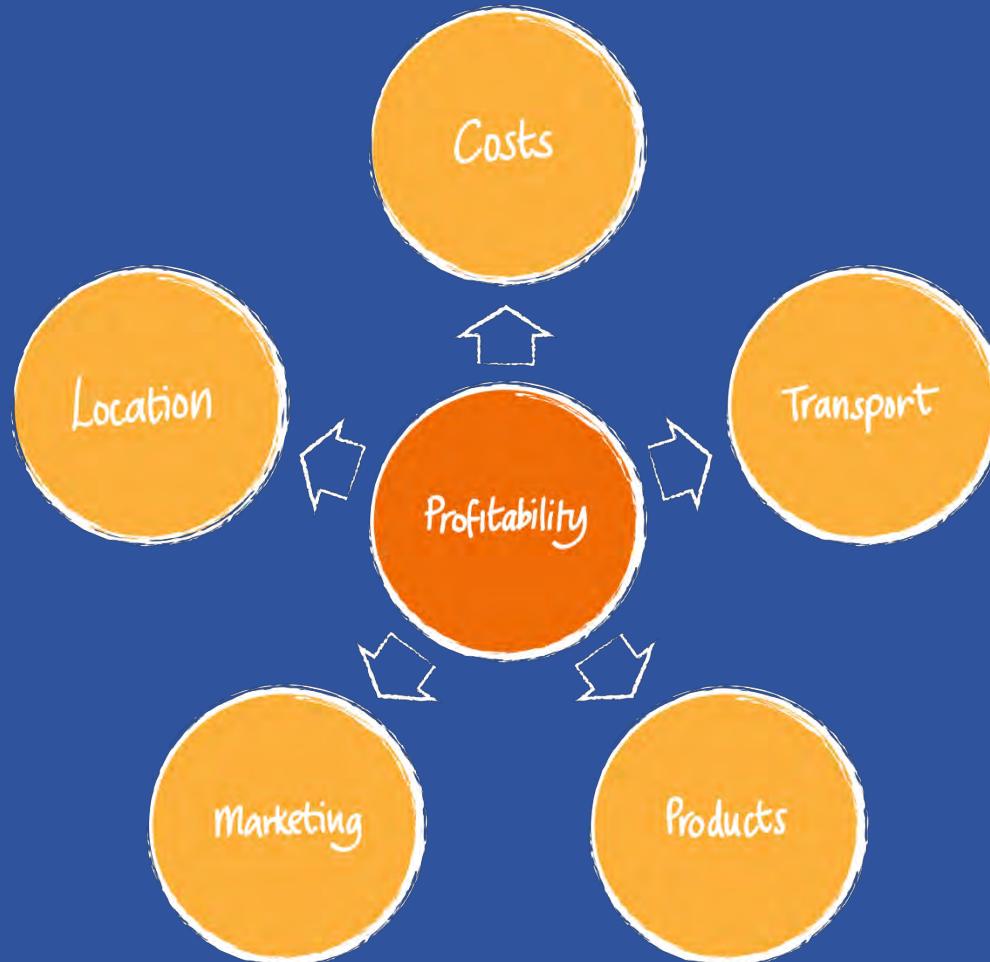
$$P(\textit{OAP, OAP}) = 0.28 * 0.28 = 0.0784$$

$$P(\textit{Both the same}) = 0.1156 + 0.1444 + 0.0784 = 0.3384$$



Areas of Improvement

We know need to start thinking about how our analyses can be used to improve profitability for the company. Here are a few areas to investigate.



Quantifying the gains

Below is the profit made by the company for the sample data you have been provided with

This Ledger makes the assumption that there are 20 bars in each box sold

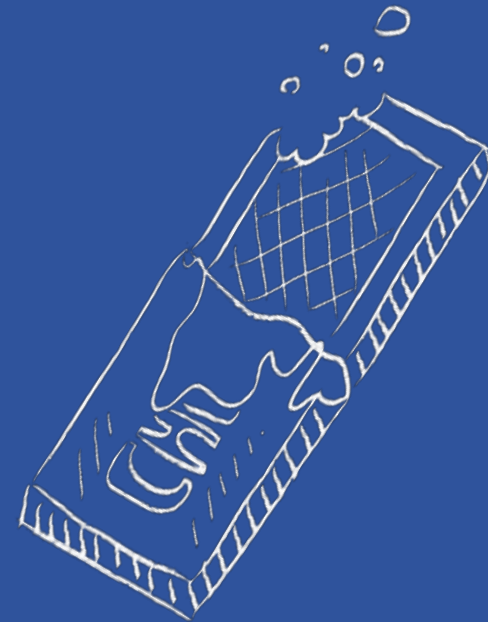
Dr.	General Ledger				Cr.
Detail	Number of Bars	€	Detail	Number of Bars	€
Large Boxes Sold	10600	€29,680.00	Large Boxes Cost	10600	€21,730.00
Medium Boxes Sold	8620	€18,102.00	Medium Boxes Cost	8620	€12,499.00
Small Boxes Sold	14700	€23,520.00	Small Boxes Cost	14700	€16,905.00
Large Bars Sold	330	€924.00	Large Bars Cost	330	€676.50
Medium Bars Sold	276	€579.60	Medium Bars Cost	276	€400.20
Small Bars Sold	491	€785.60	Small Bars Cost	491	€564.65
			Raffle	N/A	€2,600.00
Total Income		€73,591.20	Total Cost		€55,375.35
			Profit		€18,215.85

Analytics Puzzle

Three boys go into a shop to buy chocolate bars. The total cost of which is €15. They each contribute €5. The sales assistant takes the money to his supervisor who recognizes the three as regulars and asks the sales assistant to return €5 to the boys.

The sales assistant is not only poor at mathematics but is dishonest and instead of going to the trouble of splitting the €5 between the three he simply gives them €1 each and pockets the remaining €2 for himself.

Now, each of the boys effectively paid €4, the total paid is therefore €12. Add the €2 in the sales assistants pocket and this comes to €14.....Where has the other €1 gone from the original €15?

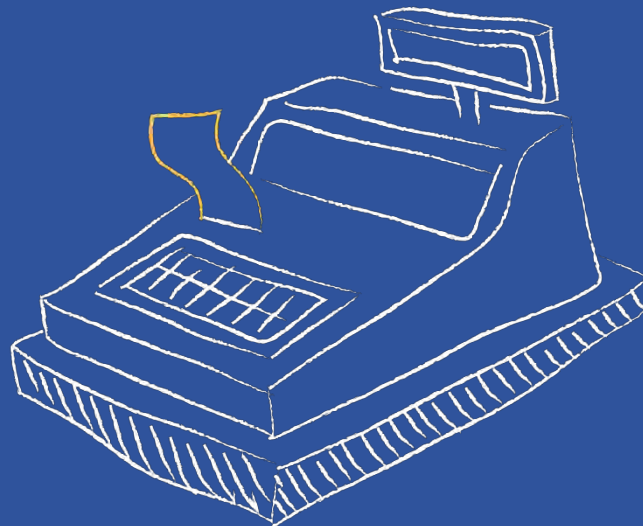
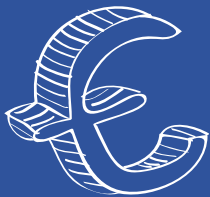


Solution on the next page!

Puzzle Solution

The payments should equal the receipts. It does not make sense to add what was paid by the boys (€12) to what was received from that payment by the waiter (€2)

Although the initial bill was €15, one of the five euro notes gets changed into five ones. The total the three boys ultimately paid is €12, as they get three ones back. So from the €12 the boys paid, the owner receives €10 and the waiter receives the €2 difference. $€15 - €3 = €10 + €2$



Next Module: New Company Strategy and the Business Report

- Forming a Strategy
- Completing a Business Report