

# ATTEMPTING A FRAMEWORK FOR THE INFERENTIAL ANALYTICS TO CREATE AN ONLINE SHOPPING DATA WAREHOUSE FOR REAL TIME PREDICTIONS

**Nishita Lamba**

*Satyawati College, University of Delhi*

---

## **ONLINE SHOPPING - CONCEPT**

Web based shopping or web based retailing is a type of electronic trade enabling customers to specifically purchase products or administrations from a merchant over the Internet without a delegate benefit.

### **SURVEY RESEARCH METHODOLOGY**

A survey has been conducted by selecting a group of people to analyse the popularity of online shopping.

### **OBJECTIVE OF THE SURVEY:**

To assess and understand the online shopping behaviour of college students further analyse the trend of online shopping in their day-to-day life

### **STATISTICAL TOOLS USED FOR THE SURVEY:**

### **SAMPLING:**

In statistics and survey methodology, sampling is concerned with the selection of a subset of individuals from within a statistical population to estimate characteristics of the whole population.

The sampling process comprises several stages:

- Defining the population of concern
- Specifying a sampling frame, a set of items or events possible to measure
- Specifying a sampling method for selecting items or events from the frame
- Determining the sample size
- Implementing the sampling plan
- Sampling and data collecting

**POPULATION:**

A population can be defined as including all people or items with the characteristic one wishes to understand.

**Population of the survey:**

The total number of students in the college.

**SAMPLING METHOD:**

Inside any of the kinds of casing recognized over, an assortment of testing techniques can be utilized, separately or in blend. These include:

1. **SIMPLE RANDOM SAMPLING:** In a clear subjective model (SRS) of a given size, each such subset of the packaging are given an equal probability. Each segment of the packaging as such has an identical probability of assurance: the packaging isn't subdivided or distributed. this points of confinement tendency and enhances examination of results.
2. **SYSTEMATIC SAMPLING:** Precise examining depends on organizing the objective populace as indicated by some requesting plan and afterward choosing components at customary interims through that arranged list. It includes an arbitrary begin and afterward continues with the determination of each kth component from that point onwards. For this situation,  $k = (\text{population measure} / \text{test estimate})$ . It is vital that the beginning stage isn't consequently the first in the rundown, however is rather arbitrarily browsed inside the first to the kth component in the rundown.
3. **STRATIFIED SAMPLING:** Where the populace grasps various particular classifications, the casing can be composed by these classifications into discrete "strata." Each stratum is then inspected as a free sub-populace, out of which singular components can be arbitrarily chosen.
4. **PROBABILITY-PROPORTIONAL-TO-SIZE ('PPS') SAMPLING:** It is one in which the determination likelihood for every component is set to be relative to its size measure, up to a limit of 1. In a basic PPS structure, these determination probabilities would then be able to be utilized as the reason for poisson inspecting.
5. **CLUSTER SAMPLING:** practical to choose respondents in gatherings ('groups'). Examining is regularly grouped by topography, or by timespans.

6. QUOTA SAMPLING: In this testing, the populace is first fragmented into fundamentally unrelated sub-gatherings, similarly as in stratified examining. For instance, a questioner might be advised to test 200 females and 300 guys between the age of 45 and 60.
7. CONVENIENCE SAMPLING: Convenience Sampling (now and then known as snatch, incidental or opportunity testing) is a sort of non-likelihood inspecting which includes the example being drawn from that piece of the populace which is promptly accessible and helpful. It might be through gathering the individual or incorporating an individual in the example when one meets them or picked by discovering them through mechanical methods, for example, the web or through telephone. .
8. LINE-INTERCEPT SAMPLING: It is a strategy for examining components in an area whereby a component is inspected if a picked line fragment, called a "transect", converges the component.
9. PANEL SAMPLING: It is the technique for first choosing a gathering of members through an arbitrary testing strategy and after that approaching that bunch for (possibly the equivalent) data a few times over some undefined time frame. Every member is met at least two time focuses; every time of information accumulation is known as a "wave".

### **SAMPLING METHOD USED IN THE SURVEY:**

#### **Convenience Sampling :**

The college students were easily accessible and reachable.

#### **SAMPLE SIZE:**

A sample of **105 students** have been chosen and analysed

#### **DATA COLLECTION:**

The data has been collected through a **Questionnaire** based on Online Shopping. A detailed questionnaire was prepared and circulated among the students and thereby the students indicated their responses as against the questions asked.

*A sample of the Questionnaire has been attached henceforth for further understanding.*

*Web based shopping questionnaire*

This survey is part of our marketing project. The objective of this survey is to assess and understand the online shopping behaviour of customers in their daily life. We request you to fill this questionnaire as part of our survey.

Please take your time and indicate your reaction by choosing the most proper decision.

Name: \_\_\_\_\_

Gender:  Male  Female

Age: \_\_\_\_\_

Occupation: \_\_\_\_\_

Family Income: (p.a.)

Less than 100000  100000 – 50000  More than 500000

**1. As a rule, you like to do your shopping of:**

ITEM/PRODUCT	Over the web	Over the Phone	In a Retail outlet
1. Groceries	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Fast food	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Cosmetics	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Books	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Clothes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**2. How often do you use the following World Wide Web Activities?**

	Never (1)	Very rarely (2)	rarely(3)	Occasionally(4)	Frequently (5)
Electronic newspaper or magazine	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Product/Service information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Shopping	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Surfing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Communication via e-mail	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**3. I have been using the web for:**

- Less than 1 year  1-2 year  2-3 year  more than 3 years

**4. How might you depict your capability on the web?**

- Novice** -just learning how to use internet.  
 **Intermediate**- I feel comfortable when using internet.  
 **Advanced**- I can use almost all internet services.

**5. Select a maximum amount (approx.) that you would spend on a solitary online buy:( Rs.)**

- Less than 500  500-1500  1500-3000  3000-5000  5000 above

**6. Select a maximum amount (approx.) you would spend per month on online purchase :(Rs.)**

- Less than 500  500-1000  1000-2000  2000 above

7. Please select your level of agreement to the following questions.				
Strongly OPPOSE(1)	Oppose(2)	Indifferent(3)	Willing(4)	Strongly Agree(5)
I think shopping on the web spares time.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is an incredible preferred standpoint to have the capacity to shop whenever of the day on the web	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is progressively hard to shop on the web.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I favor conventional shopping to web based shopping.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Shopping on the web is dangerous.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I accept web based shopping will in the end supplant conventional shopping.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I will lean toward web based shopping just if online costs are lower than real cost,	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Quite a while is required for the conveyance of items and administrations on the web	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Determination of merchandise accessible on the web is exceptionally wide,	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The portrayal of items appeared on the sites are exceptionally precise,	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The data given about the items and administrations on the web is adequate	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Internet shopping is as secure as customary shopping,	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

I hesitate to give my charge card number while shopping on the web.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Internet decreases the money costs of traditional shopping to a huge extent (parking fees etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The fact that only those with a credit card or bank account can shop on the internet is a drawback,	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would be frustrated about what to do if I am dissatisfied with a purchase made from the internet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**8.How important are the following factors in your decision to purchase goods on the web?**

Very Unimportant(1)	Unimportant(2)	Neither Important nor Unimportant(3)	Important(4)	Very Important(5)
Delivery time	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Reputation of the company	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Guarantees and Warrantees	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Privacy of the information	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Good description of goods.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Security	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Prices	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**9.How imperative are these variables in ceasing you from shopping over the web?**

Very Unimportant(1)	Unimportant(2)	Neither Important nor Unimportant(3)	Important(4)	Very Important(5)
Waiting time to get the item.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
danger of card transaction failure	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Risk of identity theft	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
trouble in returning products/items	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Risk of not receiving the correct item	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Risk of loss of privacy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Not being able to touch products	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Lack of trustworthiness of Vendors	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tough compared to Traditional Shopping	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### **PROFILE OF THE RESPONDENTS**

The respondents are students of the college. The details are as follows:

#### **OCCUPATION:**

Undergraduate students.

#### **GENDER PROFILE:**

No of Females	66
No of Males	39
TOTAL	105

**AGE GROUP:** The age group is 18-21 years

### **DATA ANALYSIS**

It is a method of evaluating, cleaning, changing, and showing data with the goal of highlighting accommodating information, proposing closures, and supporting fundamental authority. Data mining is a specific information investigation method that centers around displaying and learning disclosure for prescient as opposed to simply illustrative purposes. In measurable applications, a few people isolate information examination into elucidating insights, exploratory information investigation (EDA), and corroborative information examination (CDA).

#### **Objective:**

To analyse the product preferences of consumers while shopping on the internet.

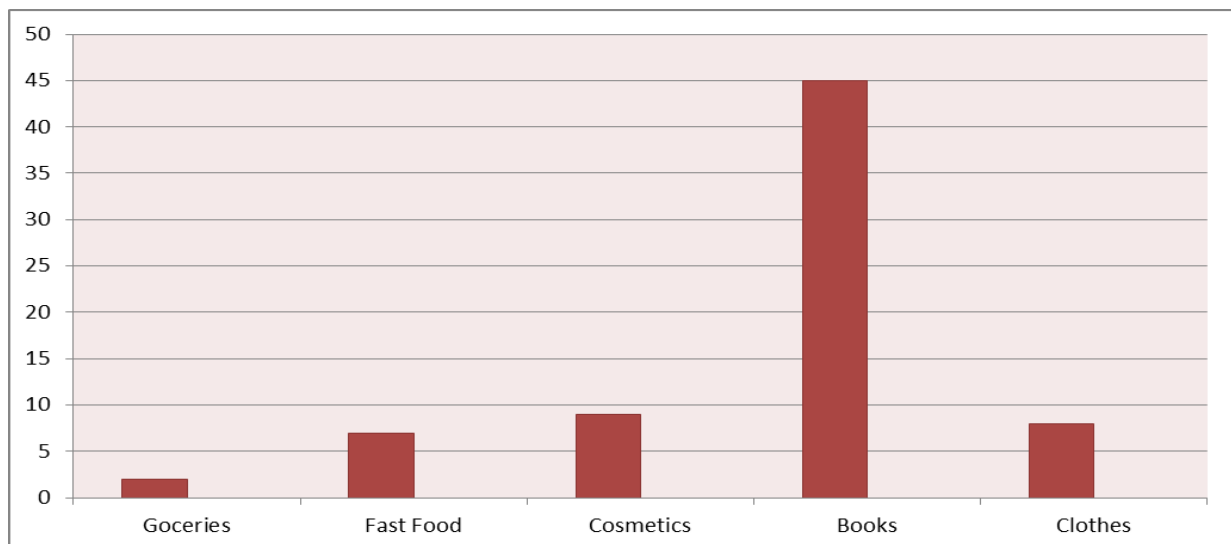
#### **Methodology:**

The students were required to indicate as to whether they shop the following products - GROCERIES, FASTFOOD, COSMETICS, BOOKS and CLOTHES-over the internet, over the phone or in a retail store.

**Findings:**

The following was found that out of the sample of **105 students**

PRODUCT	GROCERIES	FASTFOOD	COSMETICS	BOOKS	CLOTHES
No of students buying it online	2	7	9	45	8
% of students buying it online	$(2/105)*100$ =1.905%	$(7/105)*100$ =6.667%	$(9/105)*100$ =8.57%	$(45/105)*100$ =42.857%	$(8/105)*100$ 7.62%

**Graphical Representation****Implication:**

It can be clearly inferred from the graph that the product which is most preferably bought online is **Books** and least preferred are **Groceries**. The other products are also bought in a very low percentage online.



### 1. Frequency of Shopping as compared to other activities on the World Wide Web.

#### Objective:

To analyse how frequently people resort to shopping online as compared to other activities on the internet which are as follows: Electronic newspaper/magazine, product/ service information, surfing and communication via e-mail.

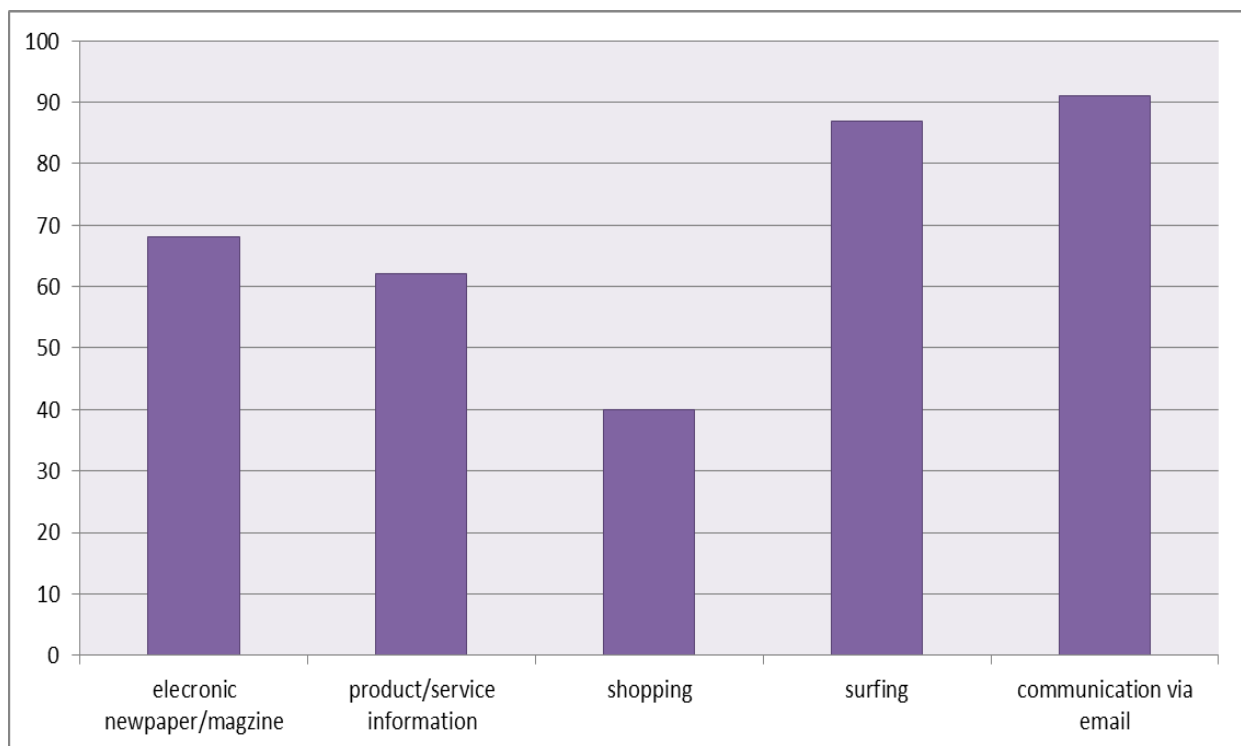
#### Methodology:

- The students were required to fill in their choices ranging from 1-5 (1 implying never using the internet for that activity and 5 implying frequent use of the internet for that particular activity ) as against the respective internet activity.
- This was an indication of how frequently the students resorted to a particular activity on the internet.
- For each activity the number of 4's and 5's were counted.
- The activity for which there were maximum 4's and 5's will be the most frequently resorted to activity on the internet.

#### Findings:

The following was found that out of the sample of 105 students

#### Graphical representation:



**Implication:**

It can be inferred from the graph that most college going students use internet for communication via email and surfing (86% and 82% respectively) whereas only 38% respondents use internet for shopping.

**2. Money spent per month as regards online shopping:****Objectives:**

- To find out the maximum amount, most students were ready to spend on online shopping transactions monthly.
- To find out how many students were ready to pay in the lowest and highest range.

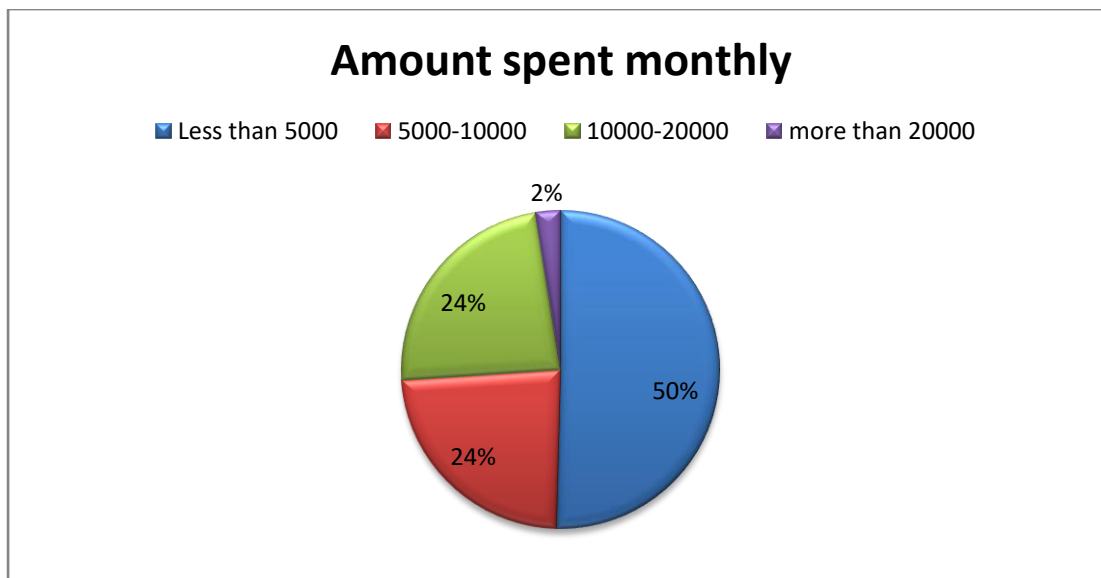
**Methodology:**

The students were asked to mark their choice in respect of the maximum amount that they would like to spend on online shopping transactions monthly.

**Findings:**

The following was found that out of the sample of **105 students** :

S.No.	Range(inRs.)	Frequency(no of students)	Percentage of students:
1	Less than 5000	46	50%
2	5000-10000	28	24%
3	10000-20000	28	24%
4	Above 20000	3	2%
Total		105	100

**Graphical representation:****Implication:**

- This implies that Most of the students( around 50%) are ready to spend less than Rs. 5000 on online shopping transactions per month.
- Around 50% students are ready to spend below 5000 Rs.
- Also, only 2% students are ready to spend more than 20000Rs.

This shows that even with an average family income of Rs.100000-500000, students still hesitate to buy expensive products on the internet and there are very less students who are ready to spend Rs.20000 on online shopping transactions per month.

**3. Factors affecting consumers' choice of shopping online:****Objective:**

To analyse the relevant importance of factors that are likely to affect the behaviour of consumers while deciding to shop online.

**Methodology:**

- The students were required to indicate their choice in the range of 1-5(**1 being very unimportant and 5 being very important accordingly**) as against the various factors affecting consumer online shopping behaviour.
- This was an indication of the level of importance that students attached to each factor.
- For each factor the number of 4's and 5's were counted.

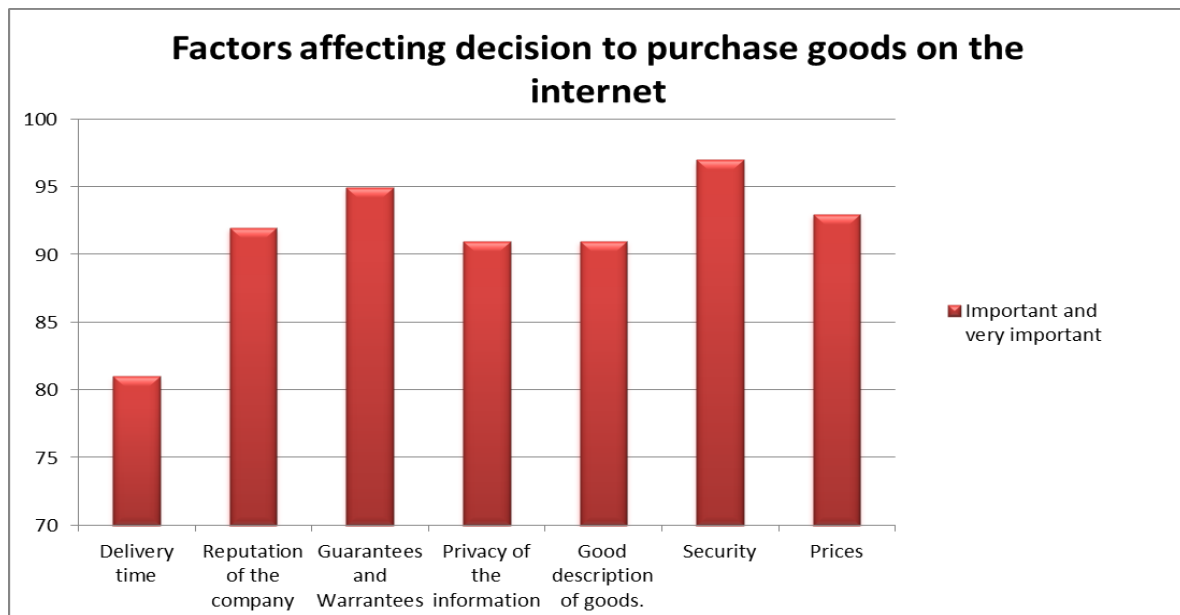
- The factor for which there were maximum number of 4's and 5's are the most important factors affecting consumer online shopping behaviour.

**Findings:**

The following was found that out of the sample of 105 students:

Factors	Delivery time	Company's Reputation	Guarantees & Warrantees	Privacy of information	Good description of goods	Security	Prices
No of students for which these factors are important	82	92	95	91	91	98	93

**Graphical Representation:**



**Implication:**

The graph clearly shows that Security is the most important factor which affects the consumers' decision to buy online, followed by prices and guarantees and warrantees.