Use of Design of Experiments in the Retail Sector

ANOUSHKA FERNANDES[1], **HEMANGI BACHHAV**[1], **ISHITA DEODHAR**[1], **KARNALI NAIK**[1], **SAYALI BORA**[1]

[1] *M.Sc. Applied Statistics, Symbiosis Statistical Institute, Pune*
**Corresponding author Email ID: anoushka.fernandes@ssi.edu.in, hemangi.bachhav@ssi.edu.in, ishita.deodhar@ssi.edu.in, karnali.naik@ssi.edu.in, sayali.bora@ssi.edu.in*

Compiled April 30, 2022

The retail sector comprises all the activities involving the sale of goods or services by a company directly to the consumer. Retail goods or services are usually purchased for personal or family use. It is a fast-paced, highly competitive, dynamic industry that has acquired steady growth year after year and employs a huge number of workers worldwide.
Use of Design of Experiments in the retail sector help improve the efficiency and quality of products and/or processes. Since experiments are time and money-consuming, the application of appropriate experimental designs helps use available resources optimally. Furthermore, the conclusions and recommendations inferred from the analysis of designed experiments are more rigorous and have greater validity than other design approaches.
In this paper, for a better understanding of the use of the design of experiments in the retail sector, 3 case studies were considered. Within them, 5 experimental designs were applied to retail stores' datasets, sales were the response variables, and conclusions of comparative analysis were drawn.

1. INTRODUCTION

The **Design of Experiments** (DOE)refers to the combination of mathematical and statistical techniques for the development of efficient, balanced, and economical experimental designs (EDs) that allow the experimenter to determine how controlled and uncontrolled factors are related to the process output(s).[2] DOE helps develop efficient, balanced, and economical experimental designs (EDs). Developed by Sir Ronald A. Fisher in the early 1920s, it empowered experimenters to systematically study the relationship between multiple input variables (factors) and key output variables (responses).

The retail sector comprises all the activities involving the sale of goods or services by a company directly to the consumer that are usually purchased for personal or family use. It has emerged as a dynamic, highly competitive, fast-paced industry and is often said to be a sort of litmus test of a country's economic level of well-being. Activities included in the retail market can be distinguished in three different macro-areas:

- **Food products** - Activities inclusive of food distribution and related products to consumers
- **Consumer goods** - Activities that sell products, most of which are reusable over time
- **Durable consumer goods** - Activities that provide consumers with longer-lasting products, e.g., household appliances, furniture, or cookware.[6]

Although physical or in-store retail is the dominant channel, forms of non-store retailing have become increasingly popular. A share of the retail sector has been carved out by online retailing or e-commerce channels in many global markets over time. The omnichannel model which aims to integrate offline and online channels in a seamless way has been adopted by many retailers in operation.

Traditional Retailers in operation include supermarkets, department stores, specialty stores, discount stores, hypermarkets, warehouse stores, etc. Retail has adapted to take advantage of the great opportunities offered by digital transformation. Internet retail which includes E-commerce and mobile retail which involves retail through smartphone applications are two new and increasingly important retailers that have entered the market.

Importance Of the Retail Sector in India:

India is the world's fifth-largest global destination in the retail space. The Indian Retail sector accounts for over 10% of the country's gross domestic product (GDP) and around 8% of the employment. Despite the coronavirus pandemic in 2020, India's retail market grew at 5.5 percent and therewith outnumbered all other global economies struggling in the face-off against the pandemic. India's retail sales were estimated to reach 1.2 trillion U.S. dollars in 2021.[5]

The country's retail industry is dominated by the unorganized or the informal sector, which is defined as businesses held by their own-account workers i.e., self-employed individuals or household members assisting them.[6]

Importance Of the Retail Sector Worldwide:

In 2021, the global retail market generated sales of over 26 trillion U.S. dollars, with a forecast to reach over 30 trillion U.S. dollars by 2024. Global retail sales were projected to amount to around 31.7 trillion U.S. dollars by 2025, up from approximately 23.74 trillion U.S. dollars in 2020. In 2021, the global in-store or brick-and-mortar retail channel generated an estimated 19.1 trillion U.S. dollars in sales. In 2021, the global in-store or brick-and-mortar retail channel generated an estimated 19.1 trillion U.S. dollars in sales. Total retail sales (online and offline retail) worldwide amounted to around 24.2 trillion U.S. dollars that year. The United States is home to the top three retail companies in the world, namely Walmart, Amazon, and Costco. [4]

Importance Of Design of Experiments in The Retail Sector:

Inferences with high confidence levels are allowed by the DOE methodology which helps optimize process management. Deployed in various types of system, process, and product design, development, and optimization, DOE is a multipurpose tool that can be used in various situations such as design for comparisons, variable screening, transfer function identification, optimization, and robust design. It is mostly used in industries during the product development phase, applied in a wide range of experiments from consumer and sensorial panel testing to stability and reliability to accelerated experiments that look at how well products maintain their shelf life. [6]

2. TERMINOLOGIES

Completely Randomized Design (CRD): It is the simplest and most flexible design. In CRD, the experimental units are allotted at random to treatments, ensuring that every unit gets the same chance of receiving every treatment. Also, in this design treatments are allocated at random to the experimental units over the entire experimental material.[1]

Model: The model for this design is given by- $y_{ij} = \mu + \tau_i + \epsilon_{ij}, i = 1,2, \dots, r_i$

where y_{ij}: response from the j^{th} unit receiving the i^{th} treatment,

μ: general mean effect

$\mathcal{T}_i$: effect due to ith treatment

ε_{ij}: error due to the random component assumed to be (i.i.d) normally distributed with mean 0 and variance σ^2

Hypotheses:

Null hypothesis, H_0: $\mathcal{T}_1=\mathcal{T}_2=\dots=\mathcal{T}_v$

Alternative Hypothesis, H_1: $\mathcal{T}1\neq\mathcal{T}2\neq\dots\neq\mathcal{T}_v$

Decision criteria:

$F_T > F_{v-1,n-v}(\alpha)$, then H_0 is rejected at α % level of significance and we conclude that treatments differ significantly
$F_T < F_{v-1, n-v}(\alpha)$, then H_0 may be accepted.

Randomized Block Design (RBD): In RBD, the experimental area is stratified into relatively homogeneous strata called blocks. The treatments are applied at random to the experimental units within each block. When all treatments appear at least once in each block, we have a completely randomized block design. [1]

Model:

$y_{ij} = \mu+\mathcal{T}_i+\beta_j+\varepsilon_{ij}$; (i= 1,2,3,...t ; j = 1,2,...,r)

where y_{ij} - response of the experimental unit receiving the i^{th} treatment in the j^{th} block;

μ- general mean effect

$\mathcal{T}_i$ -effect due to the i^{th} treatment

β_j- constant effects due to the j^{th} block

ε_{ij} - error due to the random component assumed to be (i.i.d) normally distributed with mean zero and variance σ^2.

Hypotheses:
The appropriate hypotheses for testing the equality of all the class means for treatments and blocks are:-

Null Hypothesis, $H_{0\mathcal{T}}$: $\sigma_{\mathcal{T}}^2 = 0$
$H_{0\beta}$: $\sigma_{\beta}^2 = 0$
against

Alternative Hypothesis,
$H_{1\mathcal{T}}$: $\sigma_{\mathcal{T}}^2 > 0$
$H_{1\beta}$: $\sigma_{\beta}^2 > 0$

Decision Criteria: If $F_R > F_c$, we reject $H_{0\beta}$. Similarly, for $H_{0\mathcal{T}}$

Latin Square Design (LSD): Consists of two-way stratification which helps control more variation than CRD and RBD. The whole experimental area in LSD is divided into m^2 experimental units. The number of rows, columns, and treatments is equal (m). The restriction imposed on LSD is as follows: For an experimental area divided into rows and columns, each treatment will appear only once in a row and column. [1]

Model:

$y_{ijk} = \mu + \alpha_i + \beta_j + \mathcal{T}_k + \varepsilon_{ijk}$; (i,j,k = 1,2,...,m)

where y_{ijk} - response from the unit in the ith row, jth column, and kth treatment.

μ- constant mean effect

α_i, β_j, $\mathcal{T}_k$ - constant effects due to the i^{th} row, j^{th} column, and k^{th} treatment respectively

ε_{ijk} - error due to the random component assumed to be (i.i.d) normally distributed with mean zero and variance σ^2.

Hypotheses:

Null Hypothesis, For row effects, $H_{0\alpha}$: $\alpha_1=\alpha_2=...=\alpha_m=0$,

For column effects, $H_{0\beta}$: $\beta_1=\beta_2=...=\beta_m=0$,

For treatment effects, $H_{0\mathcal{T}}$: $\mathcal{T}_1=\mathcal{T}_2=...=\mathcal{T}_m=0$

Alternative Hypothesis, For row effects, $H_{1\alpha}$: At Least two α_i's are different

For column effects, $H_{1\beta}$: At least two β_i's are different

For treatment effects, $H_{1\mathcal{T}}$: At least two $\mathcal{T}_i$'s are different

Decision criteria: If $F_R > F_c$, we reject $H_{0\alpha}$. Similarly, for $H_{0\beta}$, $H_{0\mathcal{T}}$

2^k Factorial Design: Factorial designs are most efficient to study the effects of two or more factors. In a factorial design at each complete replicate of the experiment. All possible combinations of the levels of the factors are investigated. In a 2^k factorial design effect of k factors each at two levels is being studied. In factorial experiments, the effects of several factors of variation are studied and investigated simultaneously, the treatments being all combinations of different factors under study. [2]

Let us consider the example of a 2^3 factorial experiment, we consider 3 factors A, B, C each at 2 levels, like (a0,a1),(b0,b1),(c0,c1), respectively , so that there are 23=8 treatment combinations in all. Extending the notations due to Yates for 22 experiments let the corresponding small letters a, b, and c denote the second level of each of these corresponding factors. The first level of each factor, A, B, and C is signified by the absence of the corresponding letters in treatment combinations. The 8 treatment combination in standard order are '1',a,b,ab,c,ac,bc,abc. [1]

Taguchi method:

Taguchi techniques are statistical methods established by Genichi Taguchi to enhance the quality of manufactured items. They are also known as robust design approaches. This method focuses on reducing variation without violating the target. In a dynamic response design, the quality characteristic operates along with a range of values and the goal is to improve the relationship between a signal factor and an output response.

A Taguchi design has two types of factors.

Control factors: Control factors are process or design parameters that you can control.

Noise factors: Noise factors are process or design parameters that are difficult or expensive to control during manufacturing.

Taguchi uses performance criteria that are called signal to noise (SN) ratios. A signal-to-noise ratio is a measure of robustness that can be used to identify the control factors that minimize the effect of noise on the response. [7]

Tukey's honest significance test: It is a single-step multiple comparison procedure based on the studentized range distribution and is used to find out which specific groups' means (compared with each other) are different. The test compares all possible pairs of means.

Assumptions:

1. The observations being tested are independent within and among the groups.
2. The groups associated with each mean in the test are normally distributed.
3. There is equal within-group variance across the groups associated with each mean in the test (homogeneity of variance).

Shapiro-Wilk test: The Shapiro-Wilk test, proposed in 1965, calculates a statistic that tests whether a random sample, comes from a normal distribution.

Bartlett test: Bartlett's test (Snedecor and Cochran, 1983) is used to test if samples have equal variances. Equal variances across samples are called homogeneity of variances. Some statistical tests, for example, the analysis of variance, assume that variances are equal across groups or samples. The Bartlett test can be used to verify that assumption.

3. OBJECTIVES OF THE STUDY

- To understand applications of Design of Experiment in the Retail sector.
- To apply experimental designs on appropriate datasets and observed sales for different retail products.
- To compare and contrast relevant experimental designs.

4. CASE STUDY 1

Problem Statement: To investigate the effect of shelf space on food sales. The experiment was carried out over a 6-week period using six different stores, resulting in the following data (Table 1.1) on sales of powdered coffee cream (with shelf space in parentheses.)

Aim: To compare two experimental designs i.e., RBD and LSD on the given dataset (Table 1.1)

Dataset:

		Week					
		1	**2**	**3**	**4**	**5**	**6**
Store	**1**	26(5)	14(4)	17(3)	35(1)	28(6)	21(2)
	2	34(6)	31(5)	34(4)	46(3)	37(2)	23(1)
	3	39(2)	68(6)	31(5)	51(4)	38(1)	48(3)
	4	40(3)	55(1)	39(2)	70(6)	36(4)	50(5)
	5	12(4)	15(3)	11(1)	9(2)	18(5)	15(6)
	6	16(1)	16(2)	14(6)	12(5)	15(3)	22(4)

Table. 1.1

For RBD:-

Two block designs were applied to the data:

1. The store was considered block and shelf space as treatment.
2. The week was considered block and shelf space as treatment.

Hypotheses (Stores as blocks):

H_{0B}: Mean block effects are not significantly different

H_{1B}: Mean block effects are significantly different

H_{0T}: Mean treatment effects are not significantly different

H_{1T}: Mean treatment effects are significantly different

Hypotheses (Weeks as blocks):

H_{0B}: Mean block effects are not significantly different

H_{1B}: Mean block effects are significantly different

H_{0T}: Mean treatment effects are not significantly different

H_{1T}: Mean treatment effects are significantly different

Methodology:

The Shapiro test was used to ensure that the population was normally distributed. The homogeneity of the variances was checked using Bartlett's homogeneity test. Further, Tukey's HSD test was carried out to check pairwise comparisons.

RBD was run in R once the two assumptions were met, and the following are the results.

ANOVA Table (Stores as blocks):

```
> summary(fit1)
            Df Sum Sq Mean Sq F value   Pr(>F)    
treatment    5    504   100.9   1.354    0.275    
blocks       5   6748  1349.7  18.118 1.34e-07 ***
Residuals   25   1862    74.5                     
---
```

Fig 1.1

Interpretation:

From Fig 1.1, we infer H_{0B} is rejected as the p-value is less than 0.05 and conclude that there was a significant difference in sales due to stores. There was no significant difference in sales due to shelf space.

ANOVA Table (Weeks as blocks)

```
> summary(fit2)
            Df Sum Sq Mean Sq F value Pr(>F)
treatment    5    504   100.9   0.315  0.899
blocks       5    599   119.8   0.374  0.862
Residuals   25   8012   320.5
```

Fig 1.2

Interpretation:

From Fig 1.2, we infer that H_{0B} was accepted as P-value for blocks is greater than 0.05 and concluded that there was no significant difference in sales due to the period across weeks. There was no significant difference in sales due to shelf space.

For LSD:- Stores were used as rows, weeks were used as columns, and the shelf space was the treatment effects. After the assumptions were satisfied LSD was run in R.

Hypotheses:

H_{0R}: Mean row effects are not significantly different
H_{1R}: Mean row effects are significantly different

H_{0C}: Mean column effects are not significantly different
H_{1C} Mean column effects are significantly different

H_{0T}: Mean treatment effects are not significantly different
H_{1T}: Mean treatment effects are significantly different

ANOVA table:

```
> summary(fit)
            Df Sum Sq Mean Sq F value   Pr(>F)    
treatment    5    504   100.9   1.597    0.207    
rows         5   6748  1349.7  21.367 2.14e-07 ***
col          5    599   119.8   1.897    0.140    
Residuals   20   1263    63.2                     
---
```

Fig 1.3

Interpretation:

From Fig 1.3, we infer that the p-value for the row treatment effect was less than 0.05 hence the null hypothesis was rejected for rows, and it was concluded that there was a significant difference in sales due to different Stores. The null hypothesis for treatments was not rejected, and it was concluded that there is no significant difference in sales due to shelf space.

<u>Tukey's test:</u> Since there was a significant difference in sales due to different stores, Tukey's HSD test was applied to test in which stores the sales were significantly different from one another. Table 1.1 is the R software output:

Treatments	diff	lwr	upr	p.adj
Tr05-Tr02	1.166666667	-13.25658688	15.58992021	0.9998303412
Tr04-Tr02	1.333333333	-13.08992021	15.75658688	0.9996736876
Tr01-Tr02	2.833333333	-11.58992021	17.25658688	0.9884144322
Tr03-Tr02	3.333333333	-11.08992021	17.75658688	0.976259411
Tr06-Tr02	11.33333333	-3.08992021	25.75658688	0.1802367191
Tr04-Tr05	0.1666666667	-14.25658688	14.58992021	0.9999999895
Tr01-Tr05	1.666666667	-12.75658688	16.08992021	0.9990359753
Tr03-Tr05	2.166666667	-12.25658688	16.58992021	0.9966344023
Tr06-Tr05	10.16666667	-4.256586877	24.58992021	0.2740096315
Tr01-Tr04	1.5	-12.92325354	15.92325354	0.9994209345
Tr03-Tr04	2	-12.42325354	16.42325354	0.9976937846
Tr06-Tr04	10	-4.423253543	24.42325354	0.2898182151
Tr03-Tr01	0.5	-13.92325354	14.92325354	0.999997456
Tr06-Tr01	8.5	-5.923253543	22.92325354	0.4573163605
Tr06-Tr03	8	-6.423253543	22.42325354	0.5211223941
Rows				
store6-store5	2.5	-11.92325354	16.92325354	0.9934482489
store1-store5	10.16666667	-4.256586877	24.58992021	0.2740096315
store2-store5	20.83333333	6.41007979	35.25658688	0.00237430958
store3-store5	32.5	18.07674646	46.92325354	0.000009634807387
store4-store5	35	20.57674646	49.42325354	0.000003241220712
store1-store6	7.666666667	-6.756586877	22.08992021	0.5648484206
store2-store6	18.33333333	3.91007979	32.75658688	0.008023675633
store3-store6	30	15.57674646	44.42325354	0.00002972003232
store4-store6	32.5	18.07674646	46.92325354	0.000009634807387
store2-store1	10.66666667	-3.756586877	25.08992021	0.2302264632
store3-store1	22.33333333	7.91007979	36.75658688	0.00114011053
store4-store1	24.83333333	10.41007979	39.25658688	0.0003384924087
store3-store2	11.66666667	-2.756586877	26.08992021	0.1586803374
store4-store2	14.16666667	-0.2565868768	28.58992021	0.05602631725
store4-store3	2.5	-11.92325354	16.92325354	0.9934482489
Columns				
Week1-Week3	3.5	-10.92325354	17.92325354	0.9707323439
Week5-Week3	4.333333333	-10.08992021	18.75658688	0.9298444157
Week6-Week3	5.5	-8.923253543	19.92325354	0.8323762994
Week2-Week3	8.833333333	-5.58992021	23.25658688	0.4165739032
Week4-Week3	12.83333333	-1.58992021	27.25658688	0.09929414211

Week5-Week1	0.8333333333	-13.58992021	15.25658688	0.9999677584
Week6-Week1	2	-12.42325354	16.42325354	0.9976937846
Week2-Week1	5.333333333	-9.08992021	19.75658688	0.8490646373
Week4-Week1	9.333333333	-5.08992021	23.75658688	0.3589635378
Week6-Week5	1.166666667	-13.25658688	15.58992021	0.9998303412
Week2-Week5	4.5	-9.923253543	18.92325354	0.9188027104
Week4-Week5	8.5	-5.923253543	22.92325354	0.4573163605
Week2-Week6	3.333333333	-11.08992021	17.75658688	0.976259411
Week4-Week6	7.333333333	-7.08992021	21.75658688	0.6089523552
Week4-Week2	4	-10.42325354	18.42325354	0.9489971825

Table 1.2

Interpretation - Tukey's HSD test:

From Table 1.2, we infer that the p-value for the Store 2-Store 5, Store 3-Store 5, Store 4-Store 5, Store 2-Store 6, Store 3-Store 6, Store 4-Store 6, Store 3-Store 1, Store 4-Store 1 were less than 0.05 hence we conclude that above-mentioned store pairs were significantly different from each other.

Conclusion: The mean squared error for LSD was found to be smaller as compared to RBD hence it was observed that LSD was efficient as compared to RBD

5. CASE STUDY 2

Problem statement: To test four different package designs for a new breakfast cereal. For this experiment food company chose 20 stores with approximately the same sales condition (means they have the same sales volume, price, etc.) were selected as experimental units. Five stores were randomly assigned to each of the 4 package designs. That means the first design was randomly allotted to five different stores.

Dataset:

Store Number	1	2	3	4	5
Design1	11	17	16	14	15
Design 2	12	10	15	19	11
Design 3	23	20	18	17	19
Design 4	28	26	22	33	27

Table 2.1

Here a completely randomized design (CRD) is used for the analysis of the data (Table 2.1).
As in a completely randomized design, experimental units are allotted at random to the treatments so that every unit gets the same chance of receiving every treatment. For this example, there are four different package designs treated as 4 treatments and there are twenty different stores hence total observations are 20.

Hypotheses:

H_0: The difference in mean treatment effects is not significant.

H_1: The difference in mean treatment effects is significant.

ANOVA Table:

```
> summary(anovafit)
            Df Sum Sq Mean Sq F value   Pr(>F)    
ind          3  588.2   196.1    19.8 1.24e-05 ***
Residuals   16  158.4     9.9                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Fig 2.1

Interpretation: As in the ANOVA table (Fig 2.1), the p-value is 1.24e-0.5, which is less than 0.05. So here reject the null hypothesis and conclude that there is a significant difference in different package designs for new breakfast cereal.

Tukey’s Test: Here Tukey’s test is applied. Tukey’s test is invoked when there is a need to determine if interaction among three or more variables is mutually statistically significant. In this case, Tukey’s test is used to find which package designs are significantly different.

	diff	lwr	upr	P adj
De1-De2	1.2	-4.4933576	6.893358	0.9296234
De3-De2	6.0	0.3066424	11.693358	0.0371095
De4-De2	13.8	8.1066424	19.493358	0.0000182
De3-De1	4.8	-0.8933576	10.493358	0.1147463
De4-De1	12.6	6.9066424	18.293358	0.0000537
De4-De3	7.8	2.1066424	13.493358	0.0060338

Table 2.2

As here the p-values observed in Table 2.2, for De3-De2, De4-De2, De4-De1, and De4-De3 are less than 0.05. Hence De3-De2, De4-De2, De4-De1, and De4-De3 are significantly different.

6. CASE STUDY 3

Problem statement: To investigate the effect of three different ingredients on the sale of chocolate chip cookies. An experiment was conducted in which the cookies were made using different quantities of 3 ingredients (vanilla, chocolate chips, and sugar) and the number of repeat customers (out of 100) was measured as a response. The study was conducted in eight different markets to act as replicates.

Dataset:

A: Quantity of Vanilla
B: Quantity of Chocolate Chips
C: Quantity of Sugar

A	B	C	R1	R2	R3	R4	R5	R6	R7	R8
-1	-1	-1	22	21	25	23	26	22	24	21
1	-1	-1	32	43	29	33	31	22	29	42
-1	1	-1	35	34	50	39	40	44	35	37
1	1	-1	55	47	46	45	47	43	36	59
-1	-1	1	44	41	38	41	52	45	40	51
1	-1	1	40	37	36	47	55	34	37	55
-1	1	1	60	50	54	40	50	41	37	52
1	1	1	39	41	47	46	57	44	38	46

Table 3.1

Hypotheses:

H_{01}: Main treatment effect due to treatment [a]is not significantly different.
H_{11}: Mean treatment effect due to treatment [a] is significantly different.
H_{02}: Main treatment effect due to treatment [b] is not significantly different.
H_{12}: Mean treatment effect due to treatment [b] is significantly different.
H_{03}: Main treatment effect due to treatment [c] is not significantly different
H_{13}: Mean treatment effect due to treatment [c] is significantly different
H_{04}: Mean interaction effect due to combination [a b] is not significantly different.
H_{14}: Mean interaction effect due to combination [a b] is significantly different.
H_{05}: Mean interaction effect due to combination [a c] is not significantly different.
H_{15}: Mean interaction effect due to combination [a c] is significantly different.

H_{06}: Mean interaction effect due to combination [b c] is not significantly different.
H_{16}: Mean interaction effect due to combination [b c] is significantly different.
H_{07}: Mean interaction effect due to combination [a b c] is not significantly different.
H_{17}: Mean interaction effect due to combination [a b c] is significantly different.

Methodology:

The above dataset (Table 3.1) was considered for analysis of the study. The main objective was to find which effects impact the sales of chocolate chip cookies, factorial design and the Taguchi method were used. As there are 3 factors; Vanilla, Chocolate chips, and Sugar each at two levels hence 2^3 factorial design was used. Taguchi method was used to find which factor has the most effect on the sales of the cookies. It uses the signal-to-noise ratio (S/N ratio) to identify the control factor settings that minimize the variability caused by the noise factors.

ANOVA Table:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	7	3781.94	540.28	13.47	0.000
Linear	3	2382.12	794.04	19.80	0.000
Vanilla	1	400.00	400.00	9.97	0.003
Chocolate Chips	1	280.56	280.56	7.00	0.011
Sugar	1	1701.56	1701.56	42.43	0.000
2-Way Interactions	3	1318.81	439.60	10.96	0.000
Vanilla*Chocolate Chips	1	36.00	36.00	0.90	0.347
Vanilla*Sugar	1	210.25	210.25	5.24	0.026
Chocolate Chips*Sugar	1	1072.56	1072.56	26.74	0.000
3-Way Interactions	1	81.00	81.00	2.02	0.161
Vanilla*Chocolate Chips*Sugar	1	81.00	81.00	2.02	0.161
Error	56	2246.00	40.11		
Total	63	6027.94			

Fig 3.1

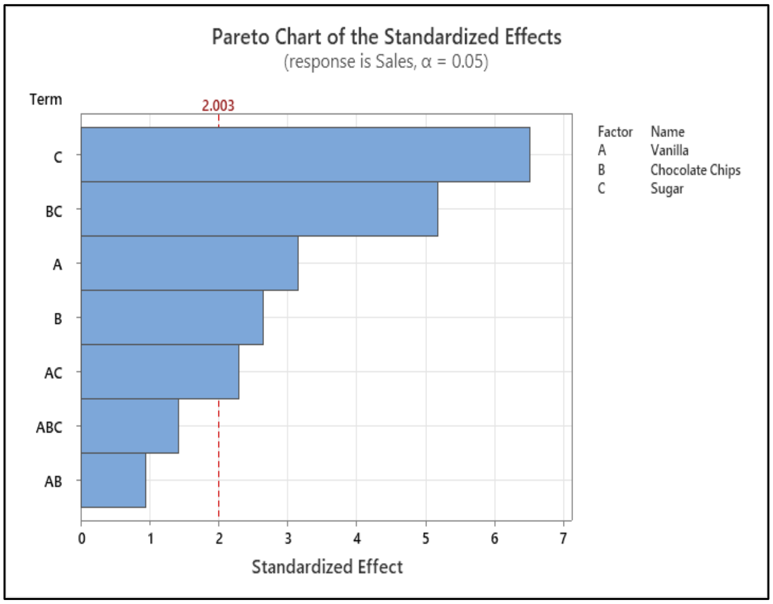

Fig 3.2

Results: The decision criteria were to reject the null hypothesis if the p-value is less than 0.05. We refer to the p-values from Fig 3.1, and make the following interpretations:

There is a significant difference in sales due to a change in the quantity of vanilla essence. There is a significant difference in sales due to a change in the quantity of chocolate chips. There is a significant difference in sales due to change in the quantity of sugar. There is a significant difference in sales due to change in the quantities of chocolate chips and sugar. There is a significant difference in sales due to change in the quantities of vanilla and sugar. No significant difference in sales due to change in the quantities of vanilla and chocolate chip cookies. No significant difference in sales due to change in the quantities of vanilla and chocolate chips cookies and sugar.

Response Table for Signal to Noise Ratios

Nominal is best (10×Log10(Ybar^2/s^2))

Level	Vanilla	Chocolate Chips	Sugar
1	17.44	17.23	18.15
2	16.63	16.85	15.92
Delta	0.80	0.38	2.23
Rank	2	3	1

Fig 3.3

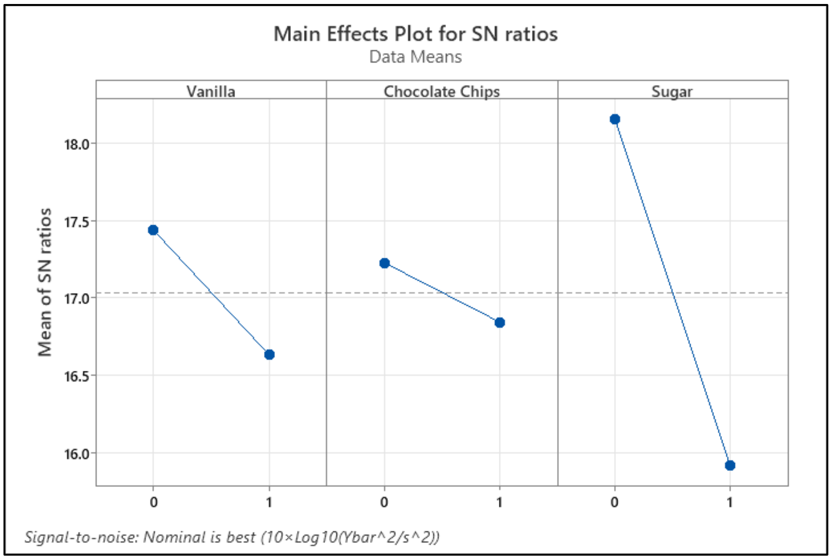

Fig 3.4

The response table for Signal to Noise Ratios gives us the delta values and ranks for all factors (Fig 3.3). The higher the delta value, the higher the rank. Here, sugar having rank 1 signifies that it causes the most effect on the response variable i.e., sales, followed by vanilla essence and chocolate chips. The same interpretation can be visualized graphically by the main effect plot for SN ratios (Fig 3.4). The more the line parallel to the x-axis the less is the effect of that factor on the sales. In general, the slope of the line gives us the magnitude of the main effect of a factor on the response variable.

7. CONCLUSION

Five relevant experimental designs were applied to three datasets that observed response variables (sales) for different retail products. In the comparative study between RBD and LSD in case study 1, LSD was concluded as the better design for the data as its MSE was lesser than that of RBD. In case study 2, CRD design was applied to the data, and we concluded that the different designs had a significant impact on the sales. In case study 3, the 2^3 factorial design and Taguchi method were applied and the factors which caused significant differences in sales of chocolate chip cookies were found. As observed, the applications of design of experiments applied to the retail sector data help retailers gain valuable insights into the measures to be taken to improve sales. Similarly, these experimental designs can be further applied to improve various other business problems and hence, provide a direction to retailers/companies to meet their business goals.

8. LIMITATIONS

- Due to restrictions on time and resources, experimental designs were applied to a small dataset. Hence, results, inferences, and/or conclusions drawn from this study may not be representative of a larger dataset.
- Relevant experimental designs other than the five designs used in this study could not be applied to the data.

9. FUTURE SCOPE

- Applications of designs of experiments in the retail industry help can further be applied to experiments that measure different response variables, e.g., yield, cost, product quality, process errors, customer satisfaction, etc.
- Furthermore, appropriate designs of experiments can be applied to study many other sectors.
- Implementation of other experimental designs, like the Response Surface Method, Central Composite Design, etc. can be applied appropriately to the many aspects of the Retail sector and a rigorous comparative study of several DOE's could be conducted.

10. REFERENCES

1. Gupta, S. C., & Kapoor, V. K. (2007). *Fundamentals of applied statistics*. New Delhi: Sultan Chand & Sons.
2. Montgomery, Douglas C. (1984). *Design and analysis of experiments*. New York: Wiley
3. Antony, Jiju & Viles, Elisabeth & Torres, Alexandre & Incerti, Taynara & Fernandes, Marcelo Machado & Cudney, Elizabeth. (2020). Design of experiments in the service industry: a critical literature review and future research directions. The TQM Journal. ahead-of-print. 10.1108/TQM-02-2020-0026.
4. Retail market worldwide - Statistics & Facts
5. Retail in India - statistics & facts
6. https://www.thebalancesmb.com/what-is-a-retailer-2890386
7. https://support.minitab.com/en-us/minitab/18/help-and-how-to/modeling-statistics/doe/supporting-topics/taguchi-designs/taguchi-designs/

11. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to Symbiosis Statistical Institute for providing opportunities for holistic development. We would like to extend our gratitude to our professor Dr. Priya Deshpande, for giving us the opportunity to explore the topic, "Design of experiments in the retail sector". The experience was much of a fruitful one. The research involved in the topic would definitely help us receive a much better understanding of the subject, which is a rather important one in the field of statistics. The experiential learning also helped us work on our team-building skills. It also helped us gain efficiency in communication, time management, and peer learning. The educational experience would prove to be beneficial to all the students.