

Design of Experiments in Search Engine Optimization; A Six Sigma Approach

Mustafa Shraim

President & Principle

SQPS LTD

Columbus, Ohio

www.sqps-ltd.com



Thairajan Rajamani

COO

Deemsys, Inc

Columbus, Ohio

www.deemsys.com



Background

- Deem Sys >> www.deemsysinc.com
- Columbus
- Services:
 - Business Intelligence
 - Consulting ERP
 - Data warehousing
 - Digital Textbook
 - E-learning courses
 - Mobile apps
 - Online tutoring
 - Project Management

Background

- SQPS >> www.sgps-ltd.com
- Columbus
- Services:
 - Lean & Quality Solutions
 - Auditing Services
 - Training & Certification
 - Coaching & Mentoring

Objective

- Increase the sale of online courses through pay-per-click (PPC) advertising



DOE & DMAIC (& PDCA)

Design of experiments process follows the Deming Cycle of Plan-Do-Check-Act (PDCA) or the Six Sigma's DMAIC process. Using the PDCA cycle, for example, the DOE process can be outlined as follows:

- **Plan:** Activities within this phase may include forming a team, brainstorming, gathering existing information, setting up the experimental design, assigning tasks for performance
- **Do:** This phase is conducting the experiment as planned and collecting data using the appropriate measures and order. For example, if randomization is necessary, then the order of experimental runs must be performed according to the selected randomization plan.
- **Check:** This refers to the analysis of the experiment using statistical and graphical methods. The analysis often includes main effects and interaction plots, analysis of variance (ANOVA), among others. Software is usually used in this step.
- **Act:** As a result of learning from the "Check" phase, a confirmation experiment should be conducted here to verify the results as well as deciding where to go next. An option might be to conduct additional or more detailed experiments, which takes us back to the *Plan* stage

Define



- Before project, online courses were marketed through
 - Organic search
 - E-mail
 - Referrals

Define – Dependent Variable

Response	Definition
Click rate	The rate (or percentage) resulting from clicking on the ad when the ad is displayed (impressions)

$$\text{Click Rate} = (\text{number of clicks}) / (\text{number of impressions})$$

Define

- This study is NOT concerned with the conversion rate
- Conversion rate depends on many other factors including features, course content, timing, and price, among others.
- A fixed weekly budget was used to run the experiment for each search engine per week

Define – Factors (Independent Variables)

Factor	Definition
<i>Search Engine</i>	Search engine used for PPC service.
<i>Ad Content</i>	The title and wording of the ad
<i>Budget</i>	A fixed weekly budget for each experimental run was used but this factor was about how to use the budget
<i>Bid Rank</i>	Rank of the ad; the higher the bid, the better the ranking. This was determined after a few trials using tools available within the search engines.

It was verified by the team that all combinations of factors can be run.

Define – Levels of Factors

ID	Factor Name	Level 1	Level 2
A	Search Engine	Search Engine 1 (SE1)	Search Engine 2 (SE2)
B	Ad Content	Conservative	Flashy
C	Budget	Evenly	To Depletion
D	Bid Rank	Top 20	Top 5

It was verified by the team that all combinations of factors can be run.

Define – Experimental Design

- A one-half fractional factorial design (2^{k-1}) was selected, where k is the number of factors.
- This means that instead of running the full experiment (2^4 or 16 runs), only one-half fraction (2^{4-1} or 8 runs) will be used.
- As a result, we would not be able to estimate all 2-way interactions

Define – Experimental Design

- As a result of this design, we would not be able to estimate all 2-way interactions
- from experience and common knowledge, the only 2-way interaction that could have an actual effect is *Budget Spending Scheme X Bid Rank* (C x D)
- The interaction (C x D) is confounded with interaction *Search Engine x Ad Content* (A x B)

Define – Experimental Design

Run	A Search Engine	B Ad Content	C Budget spent	D Bid Rank
1	SE1	Conservative	Evenly	Top 5
2	SE2	Conservative	Evenly	Top 20
3	SE1	Flashy	Evenly	Top 20
4	SE2	Flashy	Evenly	Top 5
5	SE1	Conservative	Depletion	Top 20
6	SE2	Conservative	Depletion	Top 5
7	SE1	Flashy	Depletion	Top 5
8	SE2	Flashy	Depletion	Top 20

Measure

- Each combination was run for a week
- Since 2 different search engines were used, All experimental runs were completed in 4 weeks

Measure

- Measurement of click rates were automatically provided by each of the search engines used
- Reports may be customized for any period desired

AdWords Video

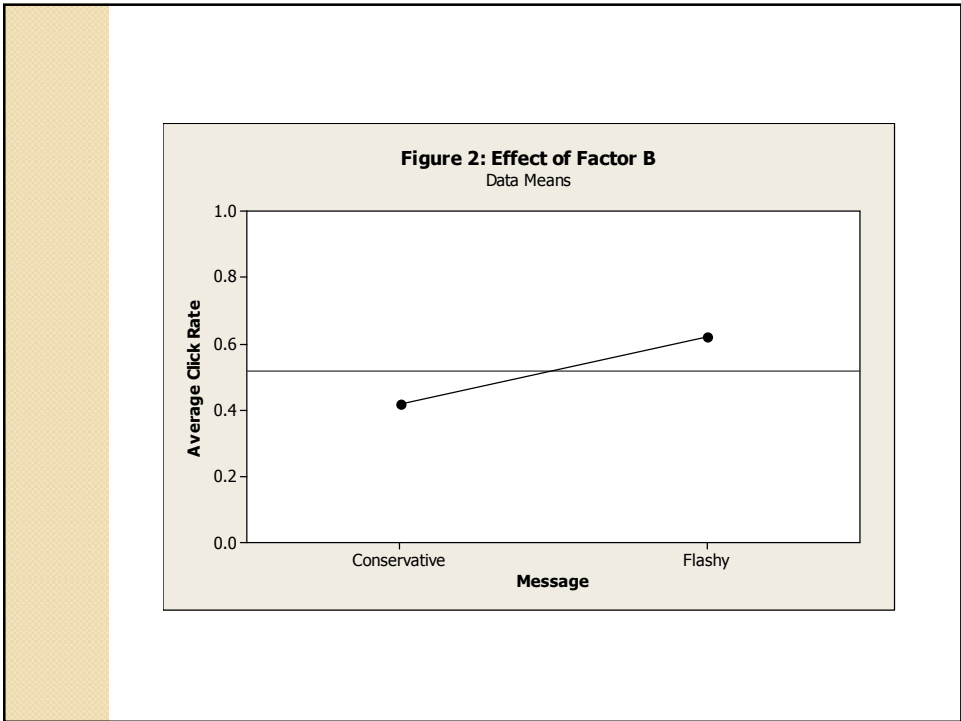
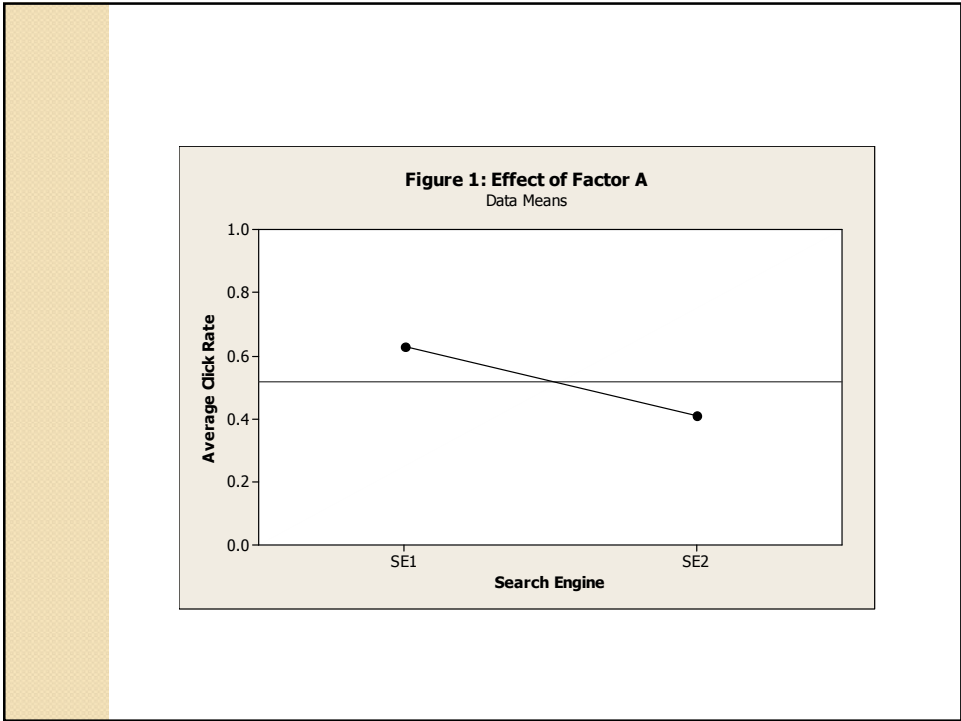
Measure – DOE Results

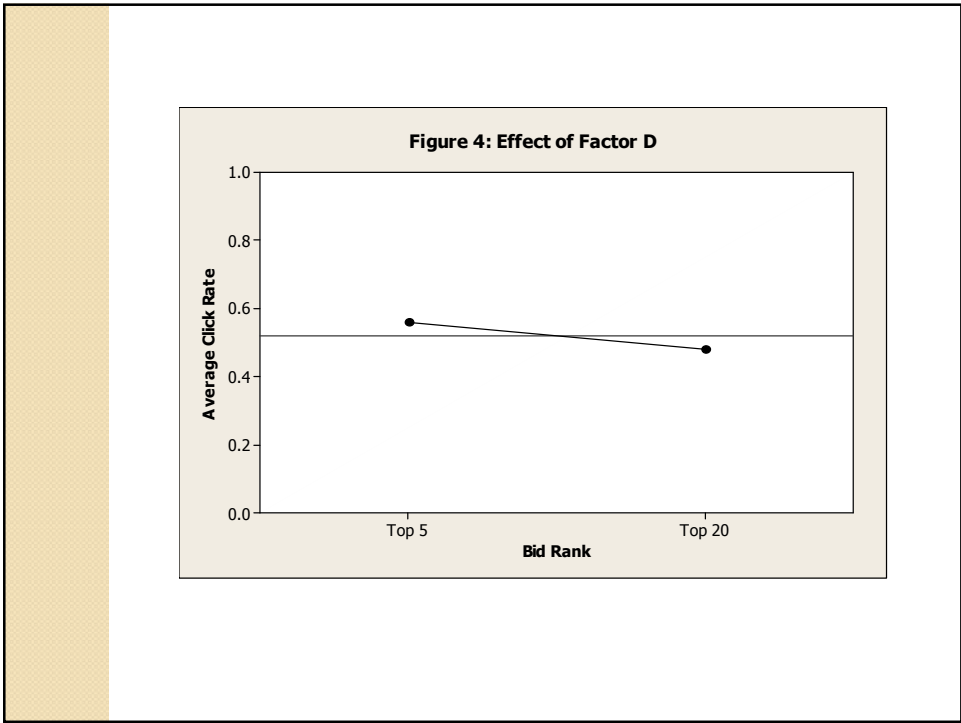
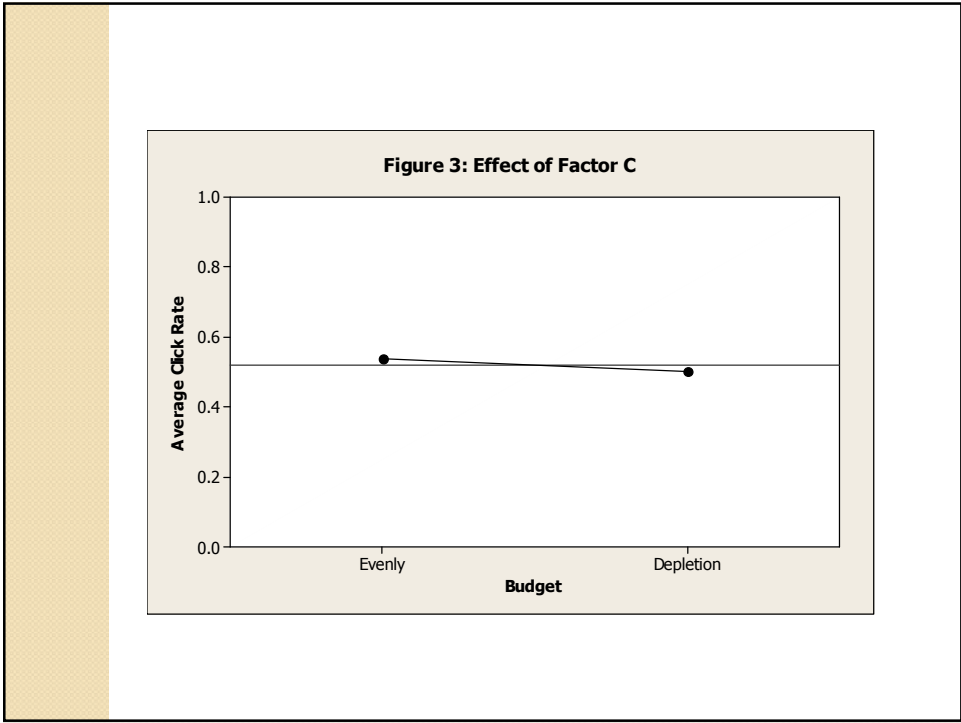
Run	A Search Engine	B Ad Content	C Budget	D Bid Rank	Click Rate
1	SE1	Conservative	Evenly	Top 5	0.63
2	SE2	Conservative	Evenly	Top 20	0.25
3	SE1	Flashy	Evenly	Top 20	0.67
4	SE2	Flashy	Evenly	Top 5	0.60
5	SE1	Conservative	Depletion	Top 20	0.50
6	SE2	Conservative	Depletion	Top 5	0.29
7	SE1	Flashy	Depletion	Top 5	0.71
8	SE2	Flashy	Depletion	Top 20	0.50

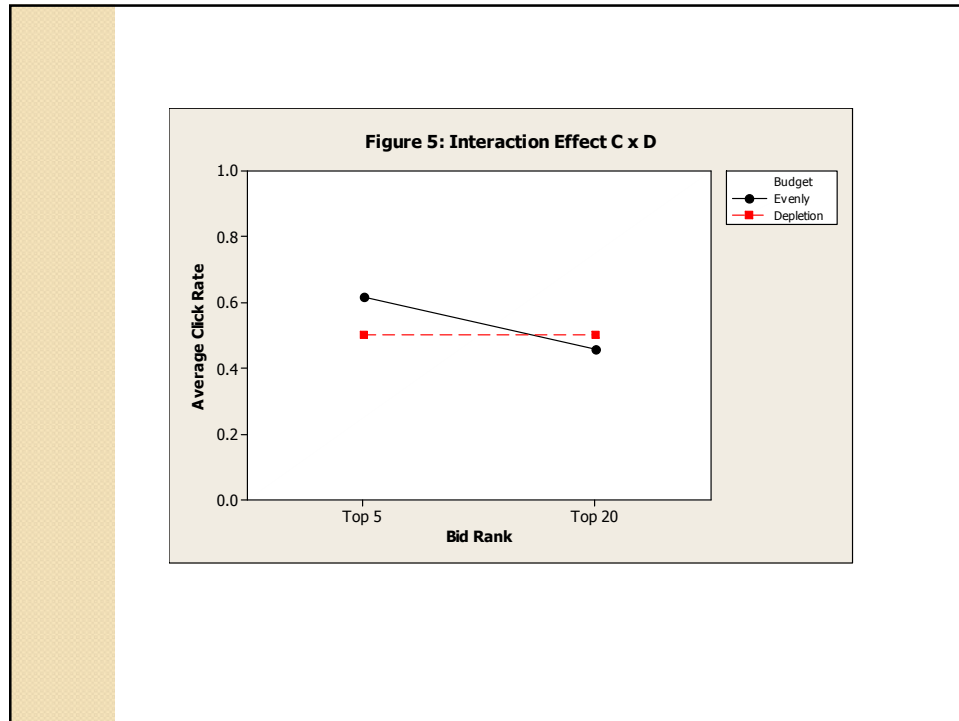
Analyze – Summary of Results

	Level 1	Level 2
A: Search Engine	63%	40%
B: Ad Content	42%	62%
C: Budget Spending Scheme	50%	54%
D: Bid Rank	56%	48%

Average Click Rates by Factor / Level







Analyze – Tests for Significant Effects

- When dealing with proportions, a typical response in marketing analysis, the effect of a factor is the difference between the average proportions corresponding to the respective levels

$$SE_{(Effect)} = \sqrt{\frac{\bar{p}(1-\bar{p})}{N_{L1}} + \frac{\bar{p}(1-\bar{p})}{N_{L2}}}$$

Analyze – Tests for Significant Effects

Null Hypothesis	$H_0: \bar{p}_{L1} = \bar{p}_{L2}$ ($\bar{p}_{L1}, \bar{p}_{L2}$ are proportions for levels 1 & 2 respectively)
Alternative Hypothesis	$H_A: \bar{p}_{L1} \neq \bar{p}_{L2}$
Test Statistic	$Z^* = \text{Effect} / SE_{(\text{Effect})}$
Significance Level	Use 5% (Z=1.96)
Decision	Reject H_0 if Effect is greater than $[SE_{(\text{Effect})} \times 1.96]$

Analyze – Tests for Significant Effects

Source	Effect $ \bar{p}_{L1} - \bar{p}_{L2} $	$SE_{(\text{Effect})}$	$SE_{(\text{Effect})} \times 1.96$	Significant at $\alpha=5\%$?
A	0.219	0.037	0.072	Yes
B	0.203	0.030	0.059	Yes
C	0.037	0.030	0.059	No
D	0.079	0.030	0.059	No
C*D	0.078	0.031	0.060	No

Analyze – Best Conditions

Factor	Setting
A (Search Engine)	SE1
B (Message)	Flashy
C (Budget)	Evenly
D (Bid Rank)	Top 5

Improve

- The best set of conditions was run for one week with the same budget that was used for each combination.
- This confirmation test resulted in a Click Rate of 0.64 (or 64%)
- About 70% of those who clicked made a purchase, suggesting that there is a demand for such an offering.

Improve

- Six Sigma projects emphasize “savings”.
- Return on investment (ROI) analysis for marketing projects are more about recovering lost opportunities rather than savings.
- Traditional marketing efforts to target potential customers may have indirectly caused the lost opportunity. In other words, a wider audience with fewer potential customers may have been targeted.
- Estimated at approximately \$185,000 for the first year alone.

Improve

Total Project Investment	\$10,250
Recovered Opportunity	\$185,000
Simple Payback Period	0.06 years or 0.66 months or 20 days
Expected recovered opportunity for 1 st year	\$174,750
Expected recovered opportunity per year thereafter	\$180,000

Control

- Weekly responses were collected to assess the performance over time.
- A control chart to be used to monitor process over time
- Traditionally, the p-chart is used to quantify the error or defect rate where a downward trend is welcome.
- Alternatively, the complement of the *Click Rate* - lost opportunity rate (LOR) may be charted.
- The p-chart with LOR data would be equivalent to the traditional p-chart where error rate is being tracked.

Control

- Another objective will be to verify that the recovered opportunity matches projections.
- To confirm this, sales data coming through the PPC service will be collected for one year after implementation.
- We should take into account any modifications we make to our settings so that we only include sales data resulting from the recommendations of this project.
- In fact, the p-chart mentioned above should give us an indication of changes that are not due to common causes.

Concluding Remarks

- As the results of this project have shown, we can utilize online metrics already available (e.g., click and conversion rates, among others) to measure the impact of marketing initiatives with minimal effort.
- Future work could focus on the number of people who actually purchase the product or service, out of those who click through. With that in mind, factors dealing with content, quality, cost, and delivery will have to be considered

Concluding Remarks

- This particular study focused on two search engines, but not all customers who might be interested in the offered course at hand are necessarily using the search engines considered in this project



Thank you!

Questions?

Mustafa Shraim

mus@shraimqps.com

614-245-0503