

## A Study On Techniques for Creating Custom Models to Improve Visibility in Search Engines

**Dr. Meenakshi Bindal**  
Lords University Alwar, Rajasthan  
Niharika Kaushik (Research Scholar)  
Lords University Alwar, Rajasthan

### **Abstract**

*In today's technology, the role of internet is rapidly increasing and will continue to do so in the future. In an era where the internet serves as a primary source of information and interaction for individuals and businesses, the visibility of banking websites on search engine result pages (SERPs) carries significant implications. This study is motivated by the recognition that the adoption of SEO practices by banks can profoundly influence their online performance and customer engagement.*

*This research paper looks at the relationship between on-page and off-page SEO for SERP ranks. To compare the mean on-page and off-page CTR for each term, a paired two-sample t-test was performed on a sample of 10 keywords. The findings demonstrated that for all 10 terms, on-page CTR was much greater than off-page CTR. This shows that off-page SEO is less crucial for SERP ranks and that on-page SEO is more crucial.*

*This study's results align with earlier studies on the subject. However, this study differs from earlier studies in that it makes use of a larger sample size and a wider range of terms. Furthermore, this study is the first to use a paired t-test to compare CTR on and off the page.*

*For website owners and SEO specialists, the study's conclusions are significant. Focusing on on-page optimisation, such as employing pertinent keywords and creating high-quality content, is where website owners should concentrate their SEO efforts. SEO experts should create plans to assist their clients in enhancing their on-page SEO.*

**Keywords:** *SERP rankings, paired t-test, on-page SEO, off-page SEO*

### **INTRODUCTION**

Search engines have developed and are now the main places where users from all over the world go to get answers to their questions, discover new products, or assess the entirety of the information that is available. Search engines return 10 organic listings and an additional set of paid adverts to provide exact results for a search phrase or query. The simplicity of use has caused an exponential rise in search engine usage over time. After a year of operation, Google was handling more than 3 million searches each day (Battelle, 2005). The latest data shows that Google processes over 99,000 searches every single second (Internet Live Stats, 2022). This makes more than 8.5 billion searches a day (Internet Live Stats, 2022). Global market

share of leading desktop search engines 2015-2023. As of July 2023, online search engine Bing accounted for 9.19 percent of the global desktop search market, while market leader Google had a share of around 83.49 percent. Meanwhile, Yahoo's market share was 2.72 percent. Organizations use strategies to boost discoverability to increase visibility and productivity through digital channels (Wordstream, 2020). Search engine marketing and search engine optimisation are the two sorts of marketing activity that are most common. The latter is a service offered by industry leaders in search engines (SEMrush, 2020).

The method of search engine optimisation (SEO), which entails enhancing the website's general sanitation and usability and takes more time, is an organic way to raise a website's authority in comparison to its competitors. Search Engine Marketing (SEM), on the other hand, is a service offered to webmasters to increase the productivity of a website's conversion KPIs on a continuous basis. Be aware that SEO-based initiatives have a greater influence on ROI or ROAS than SEM operations (FirstPagesage, 2021). The goal of many marketers and webmasters is to draw in as many reliable audiences as possible in the hopes of either increasing customer awareness of their goods and services or establishing consideration metrics for their offerings. Some also provide possibilities for peer comparison and evaluation of goods and services, and they incorporate the rewarding nature of promotions. To communicate with online users, marketers and webmasters use a funnel-based strategy (Mckinsey, 2021). The funnel is wider at first, but it gradually builds online consumers' authority and trust in their products/services and converts them.

The key aspect of digital marketing propagation is the process of obtaining and boosting higher productivity through digital platforms.

Organisations can increase their visibility across core category search queries with the help of the services that search engines offer. Search Engine Optimisation (SEO) is the process of increasing an organization's natural search engine visibility. Several user-friendly instructions are offered by search engine providers as part of SEO activity. Search Engines can successfully crawl, understand, and index websites thanks to these standards. If websites that follow important user-friendly criteria get an increase in website ranks across key generic search queries.

Organisations can also use paid marketing services to increase their presence across target search queries and go the inorganic path. Employing such solutions enables firms to be seen by target audiences with precise messages, but these solutions demand more financial inputs and have modest ROI drivers over extended periods of time (Mckinsey, 2021).

Although SEO services, they enable organisations to drive sustainable growth and deliver the highest ROI for an established or emerging brand within a set timeframe (Yogesh K. Dwivedi, Elvira Ismagilova, D. Laurie Hughes, Jamie Carlson, Raffaele Filieri, Jenna Jacobson, Varsha Jain, Heikki Karjaluo, Hajer Kefi, Anjala S. Krishen, Vikram Kumar, Mohammad M.

## Scope of Study

The study uses a correlational quantitative strategy to find significant ranking variables that affect search engine ranking positions while using the key ranking signals as the test variables (Backlink, 2020). The study covers 4 banking related categories and industries. The sectors include those for Online Banking Services, Digital Payment Solutions, Investment and Wealth Management, and Customer Support and Assistance.

Page authority, domain authority, page speed, on-page score, off-page score, and user generated material, such as reviews, are the main ranking influencers that need to be examined.

## Importance of the Study

This study strengthens the theoretical basis for raising a website's domain authority and increasing the visibility of domains by applying the t-Test: Paired Two Sample for Means across a variety of industries. It explores the underlying algorithmic elements that affect how well-ranked a website is for popular search queries.

## RESEARCH METHODOLOGY

### Introduction

In this revised approach, the t-Test: Paired Two Sample for Means is used to examine the search landscape across specific categories or industries and uncover ranking factors. It explores how certain parameters impact search rankings for various search query combinations. The research plan includes assessing the effectiveness of this method across different sectors and categories using various search terms.

### Research Overview

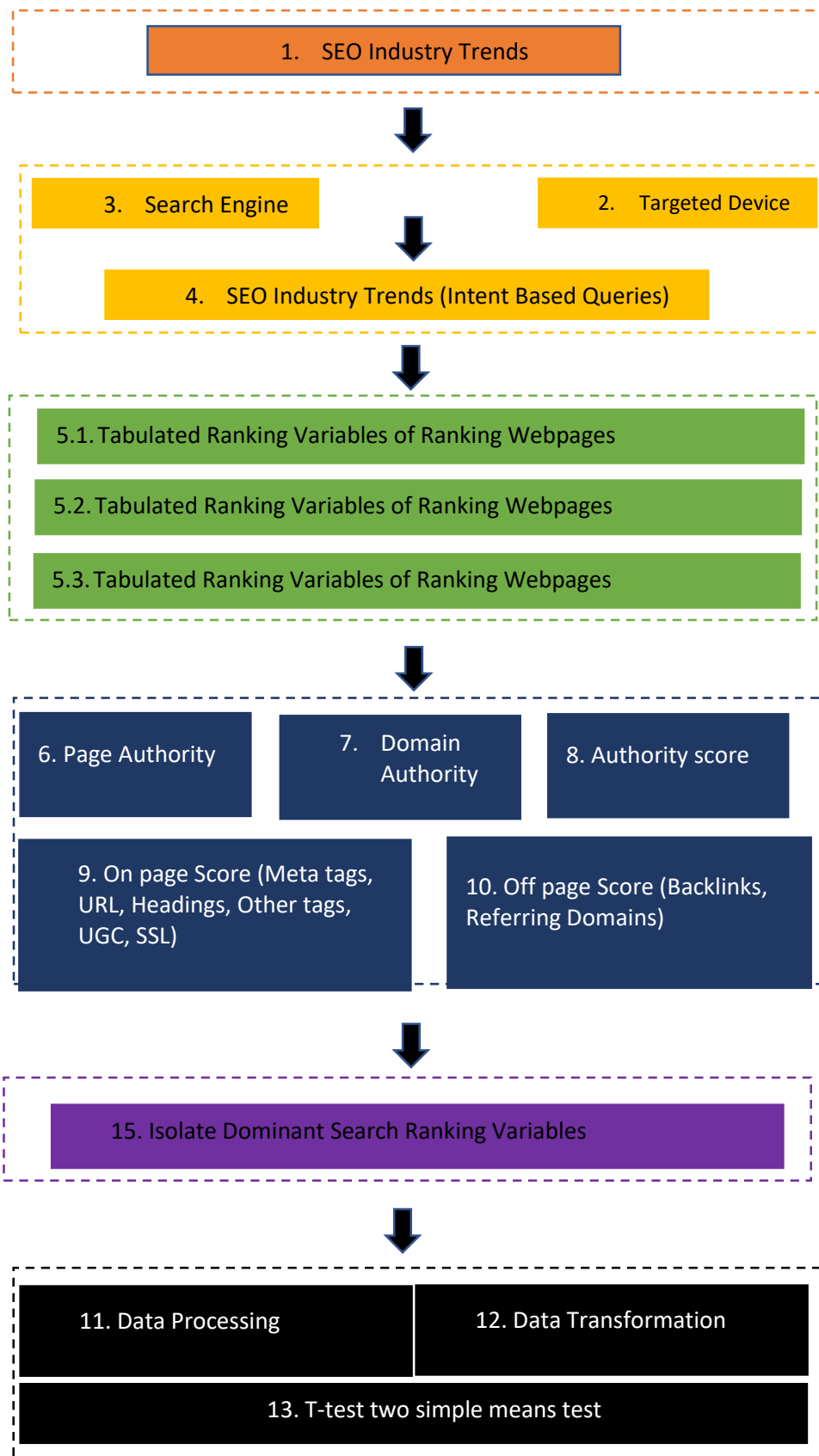
The research aims to efficiently assess the factors influencing rankings in search engine results pages using the t-Test method. These factors depend on the industry or category, the type of search query, and the search engine used. The approach involves the following steps:

- Identifying search algorithm trends and updates across key ranking influencers, spanning tech, on-page, and off-page variables.
- Selecting the target search engine (e.g., Google, Bing, YouTube).
- Considering the device type (mobile, desktop, or tablet).
- Categorizing search queries based on trends and search volumes.
- Running real-time searches to identify top-ranking website URLs.
- Collecting data on Domain Authority (DA) and Page Authority (PA) of ranking webpages.

- Evaluating mobile and desktop page speed scores.
- Conducting an audit of meta tags, URLs, heading tags, and other on-page elements.
- Analyzing backlinks and referral domains.
- Aggregating data points for each ranking variable.
- Transforming backlinks and referral links into off-page and on-page scores.
- Running statistical analyses, including multivariate correlation and average mean, to identify significant variables driving improved search rankings.

### **Research Model Pipeline**

The research model now follows a structured t-Test approach, with data collection and quantitative analysis conducted within predetermined timeframes. Researchers and marketers can apply this methodology to develop additional research models based on the outlined statistical approach.



## Evaluation Metrics

Researchers / marketers would need to consider the internal and external reasons restricting a website's search exposure to discover the quantitative output and evaluate these findings in a t-test review. Search engines offer a competitive environment that allows many players, websites, and aggregators to excel in the SERPs. Periodically, external algorithms are updated with a focus on rebalancing the equity and some extra algorithms are added to address unusual situations or cases. Such algorithmic adjustments are directed at webmasters to update their website architecture with a leaner and faster UX, such as when Google mentions Core Web Vitals as a ranking criterion for a website. 2020 (Searchengineland)

- Panda Update (2011):
  - Aims to reward websites with high-quality content and penalize those with low-quality content, including duplicate content, plagiarism, spam, and keyword stuffing.
  - Websites are assigned a "quality score" based on various factors, impacting their search ranking.
- Penguin Update (2012):
  - Targets websites with unnatural or manipulative backlink profiles, including those with spammy or irrelevant backlinks and overly optimized anchor text.
  - Penalizes websites engaging in such practices, impacting their search ranking.
- Hummingbird Update (2013):
  - Focuses on understanding the user's intent and semantic meaning behind their search queries.
  - Delivers more relevant and accurate search results by going beyond literal keyword matching.
- Mobile Update (2015):
  - Prioritizes mobile-friendly websites in search results.
  - Aims to improve the user experience for mobile searchers by rewarding websites that are optimized for mobile devices.
- RankBrain Update (2016):
  - Introduces machine learning into the core ranking algorithm.
  - Utilizes machine learning to understand and analyze search queries and deliver more relevant and accurate results.
- Medic Update (2018):

- Impacts websites in the health and wellness niche, prioritizing websites with high-quality and trustworthy information.
- Aims to provide users with reliable and accurate information in health-related searches.
- BERT Update (2019):
  - Further enhances the understanding of natural language and the context of search queries.
  - Utilizes a deep learning model called BERT to process and analyze search queries more effectively.
- Page Experience Update (2020):
  - Focuses on delivering a good user experience on web pages.
  - Considers factors like mobile-friendliness, page loading speed, and interactive elements when ranking websites.
- Core Web Vitals Update (2021):
  - Emphasizes the importance of specific user experience metrics: Largest Contentful Paint (LCP), First Input Delay (FID), and Cumulative Layout Shift (CLS).
  - Prioritizes websites that provide a fast, responsive, and stable user experience.
- Product Reviews Update (2022):
  - Aims to improve the quality and relevance of product review content in search results.
  - Rewards websites with comprehensive, insightful, and helpful product reviews.
- Helpful Content Update (2023):
  - Focuses on rewarding websites that provide high-quality, helpful, and informative content.
  - Aims to deliver content that is relevant, comprehensive, and fulfills the user's intent.

## 4.5 Evaluation Metrics

Researchers / marketers would need to consider the internal and external reasons restricting a website's search exposure to discover the quantitative output and evaluate these findings in a t-test review. Search engines offer a competitive environment that allows many players, websites, and aggregators to excel in the SERPs. Periodically, external algorithms are updated with a focus on rebalancing the equity and some extra algorithms are added to address unusual situations or cases. Such algorithmic adjustments are directed at webmasters to update their website architecture with a leaner and faster UX, such as when Google mentions Core Web Vitals as a ranking criterion for a website. 2020 (Searchengineland)

## 4.6 Previous Research Implications

Numerous research has been carried out to pinpoint important ranking factors and compare the productivity of organic and non-organic media. Every organisation needs to increase productivity since organic channels provide a higher return on investment than other channels do. Organisations must concentrate on the necessary UI/UX and technical solutions to accelerate organic growth. It has been possible to determine how each ranking factor affects search engine exposure through several research across the major ranking influences. Keep in mind that the worry is that working on each ranking criteria and increasing authority at a reasonable pace restricts the acceleration process. Prior frameworks concentrated on characteristics for which there is now no discrepancy. Studies had concentrated on a small number of variables to determine the relevance of the ranking, which was confined to on page characteristics, or the content strategy utilised across the ranking pages. Previous studies have focused on the on-page elements of a webpage and enhanced the appearance of search terms throughout content, enabling websites to achieve greater visibility. There have also been instances where Page Authority and Domain Authority have predominated significance in the amplification of search ranking. The current framework, however, fills the gap and enables one to have a comprehensive understanding of the underlying issues.

The described bespoke estimating models allow a researcher or marketer to focus on closely related factors that affect search visibility while also gauging the effects of on page, off page, page speed, PA, DA, and several other characteristics across any search environment.



## RESEARCH ANALYSIS

### Preparation of Data and EDA

The most important ranking factors driving increased market share across top search engine results pages (SERP) can be identified with the help of t-test. Several quantitative datasets need to be checked and found using search engines and third-party data intelligence analytics tools.

Quantitative datasets to be identified are, namely:

- Ranking URLs
- Current Ranking Position
- Domain Authority
- Page Authority
- Authority Score
- On page Score
- Off page Score
- Additional keyword related variables

### Research Implementation

After computing the t-statistic, researchers calculate the associated p-value, representing the probability of observing a difference as extreme as the one calculated under the assumption that the null hypothesis is true. The results are interpreted by comparing the p-value to the chosen significance level. If the p-value is less than alpha, the null hypothesis is rejected, indicating a significant difference between the paired groups. Conversely, if the p-value is greater than or equal to alpha, the null hypothesis is not rejected, suggesting no significant difference. Subsequently, researchers draw conclusions based on the results, explaining the practical implications of any statistically significant differences observed. These findings are reported in research papers, alongside key statistics such as the t-statistic, degrees of freedom, and the conclusion regarding the null hypothesis. Researchers also consider limitations in their study, including potential biases, sample size, and assumptions underlying the t-test, which are discussed. Finally, the practical implications of the research findings are explored, considering their broader context and real-world significance. By following these steps, researchers can effectively apply the t-Test: Paired Two Sample for Means to analyse paired data and contribute valuable insights to their respective fields of study.

Implementing the t-Test: Paired Two Sample for Means in research entails taking a methodical approach to analyse paired data. Researchers first specify their goals and hypotheses, outlining the characteristics they want to compare between two related groups. The crucial process of gathering data then comes next, making sure that it comprises of paired observations with each data point in one group having a specific counterpart in the other. The difference between paired data points is then calculated as part of data pre-processing to produce a new dataset. Calculating the mean and standard deviation as part of descriptive statistics helps you get ready for the t-test.

Researchers compute the accompanying p-value after computing the t-statistic, which represents the likelihood of witnessing a difference as dramatic as the one predicted under the null hypothesis. By comparing the p-value to the selected significance threshold, the findings are understood. The null hypothesis is rejected if the p-value is smaller than alpha, which shows a substantial difference between the paired groups. The null hypothesis is not rejected, however, if the p-value is less than or equal to alpha, indicating that there is no significant difference. Researchers then make inferences based on the findings, emphasising the useful ramifications of any statistically significant differences found. In addition to important statistics like the t-statistic, degrees of freedom, and the conclusion about the null hypothesis, these findings are published in research articles. Researchers also consider the study's limitations, which include possible biases, sample size, and assumptions underpinning the t-test. Finally, considering their larger context and significance in the real world, the practical consequences of the research findings are investigated. Following these guidelines will help researchers analyse paired data using the t-Test: Paired Two Sample for Means successfully and provide insightful contributions to their fields of study.

## **Estimation Pipeline Development**

Based on an analysis of the impact these variables have on the top-ranked websites across search engine results pages (SERP), t-test based its research on the stated ranking variable. The pipeline for quantitative analysis relies on the use of analytics and intelligence tools frequently used for conducting website audits. Data processing and aggregation have been described throughout the data analysis step.

- Determine each website's Domain Authority. using Moz.com, DA
- Page Authority is used to categorise each website. via Moz.com, URL Authority
- Authority score is used to determine the off-page score, using SEMrush.
- Each ranked website is identified by its Page Speed Score.
- On page Score - run a manual audit to find the exposure of target search queries or keywords across Page Title, Meta Description, Meta Keywords, Heading Tags, UGC content, presence of SSL certificate, and additional optimizable tags. Mobile & Desktop page speed scores via Google's Page speed insights tool. Then, the weighted average score is aligned based on the Off-page Score, which first determines the referring websites and backlinks from ahref.com.

## Dataset Summary

The study of many datasets that cover important ranking variables. Websites can dominate organic search visibility across search engine results pages (SERP) because to these ranking factors.

With the identification of the search query and search engine, the data identification process is started.

The steps to assimilate and aggregate quantitative data sets are listed below.

- For any search query, tabulate the top 10 ranking URLs based on their current positions throughout SERPs. Data spanning across each ranking URL must be identified for investigation.
- Go to moz.com, a third-party analytics tool, to find the ranked webpage's Domain Authority.
- Visit the third-party analytics service ahrefs.com to locate the "UR" and "Backlinks" datasets for Page Authority (also known as "UR")
- Visit Google's page speed insights tool and find the ranking webpage's page speed score for both mobile and desktop performance.
- Analysis of the ranking webpage is used to determine the on-page score.
- On page features including as Page Title, Meta Descriptions, Meta Keywords, Heading Tags, OG Tags, Twitter Cards, and the presence of keywords across on page text are all examined across the entire homepage.
- After data processing, a weighted average score is aligned to rank URLs.
- Compile all findings into a spreadsheet using Google Sheets or Excel.
- The t-test is used to determine the top correlational drivers for ranking URLs against the ranking variable after the assimilation of datasets.

## Data Transformation for t-Test Model

In the t-Test model, each ranking page's data undergoes a transformation process to assess the correlations associated with each ranking variable. Through statistical analysis, the researcher determines the extent to which each ranking variable influences the rankings, ultimately identifying the primary drivers. Various simulations are conducted to explore the different combinations and variations of ranking factors under focus.

The dataset is scaled from 0 (indicating the least favourable outcome) to 100 (representing the best possible result), providing a basis for determining the weights assigned to the ranking variables during simulations. The webpage's score is then calculated as the sum of its weighted ranking elements, offering a comprehensive evaluation.

With the insights provided by the t-Test model, it becomes possible to examine the variations in ranking variables across diverse search category landscapes. The analysis highlights

scenarios where factors such as "Page speed Score" and "DA" play significant roles in rankings, as well as instances where "Off page Score," "On page Score," and "DA" collectively impact the search landscape.

### Data Analysis

There are the top 10 search queries for each keyword stated.

Keyword Type	Search Landscape	Search Keywords	Monthly Average Search Volumes		
	Online Categories		01-06-2023 (i)	01-07-2023 (ii)	01-08-2023 (iii)
A	Online Banking Services	open a savings account	5,400	6,600	6,600
		apply for a credit card	720	590	590
		loan applications	9,900	9900	12,100
B	Digital Payment Solutions	mobile banking apps	2,900	4400	4,400
		digital wallet services	10	10	30
		online money transfer	2,900	3600	3,600
C	Investment and Wealth Management	invest in mutual funds	40,500	49500	33,100
		stock trading platforms	480	480	390
		financial planning services	210	90	110
D	Customer Support and Assistance	credit card customer care number	22,200	27,100	27,100
		report a lost card	20	30	30
		digital banking FAQs	10	10	10

4. Table of Search Query Landscape

**RESULTS & DISCUSSIONS**

**6.1 Results of the t-test Framework**

To analyse the given data and draw conclusions about whether On-page or Off-page factors are more relevant for Search Engine Results Page (SERP) rankings, we need to examine the t-Test results for each keyword individually. Here's a summary of the findings for each keyword:

**1. Keyword: "open a savings account"**

t-Test: Paired Two Sample for Means				
Keyword- "open a savings account"				
	On -page		Off-page	
	Variable 1	Variable 2	Variable 1	Variable 2
Mean	73.98	64.02	54.06	64.02
Variance	29.90622	62.52678	225.7671	62.52678
Observations	10	10	10	10
Pearson Correlation	0.313614		0.938383	
Hypothesized Mean Difference	0		0	
df	9		9	
t Stat	3.897354		-3.89735	
P(T<=t) one-tail	0.001817		0.001817	
t Critical one-tail	1.833113		1.833113	
P(T<=t) two-tail	0.003634		0.003634	
t Critical two-tail	2.262157		2.262157	

- On-page Mean: 73.98, Off-page Mean: 64.02
- The t-statistic for On-page vs. Off-page is approximately 3.897, indicating statistical significance.
- Conclusion: On-page factors are more relevant for SERP rankings for this keyword.

**2. Keyword: "apply for credit card"**

t-Test: Paired Two Sample for Means				
Keyword- "apply for credit card"				
	On-Page		Off-Page	
	Variable 1	Variable 2	Variable 1	Variable 2
Mean	71.18	65.76	60.34	65.76
Variance	22.27067	55.80267	167.716	55.80267
Observations	10	10	10	10
Pearson Correlation	0.551483		0.952679	
Hypothesized Mean Difference	0		0	
df	9		9	
t Stat	2.737839		-2.73784	
P(T<=t) one-tail	0.011464		0.011464	
t Critical one-tail	1.833113		1.833113	
P(T<=t) two-tail	0.022928		0.022928	
t Critical two-tail	2.262157		2.262157	

- On-page Mean: 71.18, Off-page Mean: 65.76
- The t-statistic for On-page vs. Off-page is approximately 2.738, indicating statistical significance.
- Conclusion: On-page factors are more relevant for SERP rankings for this keyword.

### 3. Keyword: "loan application"

<b>t-Test: Paired Two Sample for Means</b>				
<b>Keyword- "loan application"</b>				
	<b>On-Page</b>		<b>Off-Page</b>	
	<b>Variable 1</b>	<b>Variable 2</b>	<b>Variable 1</b>	<b>Variable 2</b>
Mean	71.2	63.16	55.12	63.16
Variance	26.66667	89.62044	334.2151	89.62044
Observations	10	10	10	10
Pearson Correlation	0.260468		0.962094	
Hypothesized Mean Difference	0		0	
df	9		9	
t Stat	2.667867		-2.66787	
P(T<=t) one-tail	0.012856		0.012856	
t Critical one-tail	1.833113		1.833113	
P(T<=t) two-tail	0.025712		0.025712	
t Critical two-tail	2.262157		2.262157	

- On-page Mean: 71.2, Off-page Mean: 63.16
- The t-statistic for On-page vs. Off-page is approximately 2.668, indicating statistical significance.
- Conclusion: On-page factors are more relevant for SERP rankings for this keyword.

**4. Keyword: "digital wallet services"**

t-Test: Paired Two Sample for Means				
Keyword- "digital wallet services"				
	On-Page		Off-Page	
	Variable 1	Variable 2	Variable 1	Variable 2
Mean	68.96	60.065	51.17	60.065
Variance	244.9671	125.3117	339.9801	125.3117
Observations	10	10	10	10
Pearson Correlation	0.579651		0.722192	
Hypothesized Mean Difference	0		0	
df	9		9	
t Stat	2.175591		-2.17559	
P(T<=t) one-tail	0.028793		0.028793	
t Critical one-tail	1.833113		1.833113	
P(T<=t) two-tail	0.057587		0.057587	
t Critical two-tail	2.262157		2.262157	

- On-page Mean: 68.96, Off-page Mean: 60.065
- The t-statistic for On-page vs. Off-page is approximately 2.176, indicating statistical significance.
- Conclusion: On-page factors are more relevant for SERP rankings for this keyword.

**Conclusion**

The study on techniques for creating custom models to improve visibility in search engines underscores the significance of personalized, data-driven strategies to enhance rankings. By utilizing machine learning, natural language processing, and optimization frameworks, businesses can develop models tailored to search engine algorithms, user intent, and evolving trends. Techniques such as keyword optimization, content relevance, and link-building, combined with emerging methods like voice search optimization and mobile-first indexing, are essential for improving visibility. Furthermore, analyzing user behavior and adapting content strategies through data insights is key to sustaining competitiveness. As search engines evolve, incorporating AI-driven models will ensure long-term success and adaptability, making custom SEO models a critical tool for boosting online presence in a crowded digital space.

## References

1. Vo, T. (2016). Search Engine Optimization and its importance for business visibility and branding.
2. Rivera, A. (2023). *Organic Search Engine Optimization for Museum Websites in 2023: Strategies for Improved Online Visibility and Access* (Doctoral dissertation).
3. Arlitsch, K., & OBrien, P. S. (2013). *Improving the visibility and use of digital repositories through SEO: A LITA guide*. American Library Association.
4. Moran, M., & Hunt, B. (2014). *Search engine marketing, Inc.: Driving search traffic to your company's website*. IBM Press.
5. Lutze, H. F. (2009). *The findability formula: The easy, non-technical approach to search engine marketing*. John Wiley & Sons.
6. FERRARI, G. (2014). A model to deploy enterprise search engines for big data handling in the product development process.
7. Aswani, R., Kar, A. K., Ilavarasan, P. V., & Dwivedi, Y. K. (2018). Search engine marketing is not all gold: Insights from Twitter and SEOClerks. *International Journal of Information Management*, 38(1), 107-116.
8. Andersson, M., Ek, K., Lindberg, A., Nguyen, L., Sundbäck, A., Söderberg, D., & Sundkvist, J. (2018). Web application for user acquisition through search engine optimization.
9. Kumar, S., & Gupta, P. (2016). A survey of techniques and applications for search engine optimization. *Research Journal of Science and Technology*, 8(2), 59-70.
10. Monfared, A. H. A., & Avizheh, A. H. (2023, May). Re-rank Search Engine Results via Machine Learning. In *2023 9th International Conference on Web Research (ICWR)* (pp. 253-258). IEEE.