SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

FACULTY OF SCIENCE AND HUMANITIES

DEPARTMENT OF COMPUTER APPLICATIONS

SRM Nagar, Kattankulathur – 603 203

<u>INTERNSHIP REPORT</u>

STUDENT NAME **: ARUNKUMAR A**

REGISTER NUMBER **: RA2232014010056**

CLASS **: II MSC ADS   Section: "A"**

YEAR & SEMESTER **: II YEAR – III SEMESTER**

SUBJECT CODE **: PAD21P01L**

SUBJECT TITLE **: INTERNSHIP**

OCTOBER 2023

I

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

FACULTY OF SCIENCE AND HUMANITIES

DEPARTMENT OF COMPUTER APPLICATIONS

SRM Nagar, Kattankulathur – 603 203

CERTIFICATE

Certified to be the bonafide report of internship work done by

*_____ Register No. _____ of **MSc Applied Data***

***Science** Degree course for **PAD21P01L – INTERNSHIP** during the academic*

year 2023-2024.

                                                                              Head of the Department

Submitted for Semester Practical Examination held on _____.

Internal Examiner   *External Examiner*

DATA SCIENCE UPSKILLING PROGRAM

NAME   : **ARUNKUMAR A**

REGISTER NUMBER   : **RA2232014010056**

INTERNSHIP DURATION   : **5ᵀᴴ APRIL TO 5ᵀᴴ JUNE 2023**

COMPANY NAME   : **SKILL VERTEX**

TABLE OF CONTENTS

I EXECUTIVE SUMMARY

This executive summary encapsulates the enriching experience of my Data Science Internship at Skillvertex, which took place from 5th April 2023 to 5th June 2023. The primary objectives of this internship were to develop a comprehensive understanding of data science techniques, enhance data analytics skills, and apply statistical and hypothesis testing methods to real-world data scenarios.

Throughout the internship, I actively engaged in rigorous training sessions and hands-on exercises, which provided a solid foundation in data science fundamentals. These activities facilitated a profound understanding of data collection, cleaning, analysis, and the practical application of statistical methodologies.

Key accomplishments during this internship included the design and implementation of data pipelines, leveraging data science tools and libraries for effective data processing. I also collaborated with the data science team to conduct in-depth statistical analysis and hypothesis testing, resulting in data-driven decisions that significantly influenced project outcomes.

The expectations for this internship were not only met but exceeded. I had the privilege of contributing meaningfully to data-driven projects, aligning with Skillvertex's data strategy. The mentorship and guidance of experienced professionals were invaluable in my professional growth, and I had the opportunity to apply my academic knowledge to real-world data challenges.

This internship has equipped me with a diverse skill set, including data collection, data cleaning, and data analysis, as well as proficiency in using data science tools and libraries. Furthermore, I have developed the capability to conduct statistical analysis and hypothesis testing, enhancing problem-solving, critical thinking, and communication skills. These newfound skills are pivotal in my academic and professional journey, allowing me to tackle complex data-related challenges effectively and make informed, data-driven decisions.

II INTRODUCTION

Introduction:

The purpose of this report is to provide an insight into my Data Science Internship experience at Skill vertex, spanning from 5th February 2023 to 5th April 2023. This introduction will delineate the objectives of the internship, introduce Skill vertex as the host organization, and shed light on its history, mission, and core activities.

Purpose of the Internship:

The primary purpose of this internship was to gain practical experience in the field of data science and analytics. It aimed to bridge the gap between theoretical knowledge and its application in real-world data scenarios. The overarching objectives included:

- Comprehensive Understanding of Data Science Techniques: To acquire a profound understanding of data science techniques, tools, and methodologies, and their practical applications.
- Improving Data Analytics Skills: To enhance data analytics skills, encompassing data collection, cleaning, visualization, and interpretation, with the goal of extracting meaningful insights from data.
- Application of Statistical and Hypothesis Testing Methods: To apply statistical analysis and hypothesis testing methods to real-world data scenarios, enabling data-driven decision-making and hypothesis evaluation.

About Skill vertex:

Skill vertex is a distinguished institution specializing in data science education and training. Located at the forefront of data science education, Skill vertex is renowned for its commitment to empowering individuals with data-driven knowledge and skills

Company Overview:

History: Skill vertex has a rich history of educating and training data science professionals since its inception. Founded in 2021, it has continually evolved to stay current with the latest advancements in data science and analytics.

Mission: The mission of Skill vertex is to provide quality education and training in the field of data science, equipping individuals with the skills and knowledge required to excel in data-driven professions.

Products and Services: Skill vertex offers a wide range of products and services, including data science courses, workshops, certifications, and online resources. These offerings cater to students, professionals, and organizations seeking to harness the power of data.

Skill vertex's expertise in data science education and training makes it an ideal platform for practical internships, aligning with its mission to prepare individuals for real-world data challenges.

The following sections of this report will delve into the details of my internship experience, including the activities, projects, skills acquired, and the impact of this experience on my personal and professional development.

III OBJECTIVES AND EXPECTATIONS

OBJECTIVES:

Comprehensive Understanding of Data Science Techniques:

- The primary objective was to develop a comprehensive understanding of data science techniques, methodologies, and tools.

- This encompassed gaining practical knowledge in data collection, cleaning, analysis, and the application of data science concepts to real-world problems.

Improving Data Analytics Skills:

- Another key objective was to enhance data analytics skills.

- This included improving proficiency in data cleaning, data visualization, and the interpretation of data to derive meaningful insights.

Application of Statistical and Hypothesis Testing Methods:

- A crucial objective was to apply statistical analysis and hypothesis testing methods to real-world data scenarios.

- This involved conducting statistical analyses, evaluating hypotheses, and making data-driven decisions based on the results.

EXPECTATIONS:

Meaningful Contributions to Data-Driven Projects:

- The expectation was to actively contribute to data-driven projects that align with Skill vertex's data strategy.

- This involved applying data science techniques to real-world problems and providing valuable insights to support data-driven decisions.

Mentorship and Professional Growth:

- Expectations included benefiting from the mentorship and guidance of experienced data science professionals.

- The internship aimed to provide a platform for personal and professional development, learning from experts in the field.

IV DESCRIPTION OF DUTIES AND ACCOMPLISHMENTS

During my internship at Skill vertex, I undertook a variety of tasks and responsibilities, which contributed to a rich learning experience and meaningful contributions to the organization. These included:

Data Collection and Preprocessing:

- I was responsible for collecting and preprocessing data from various sources. This involved data cleaning, transformation, and ensuring data quality for further analysis.

- Accomplishment: Successfully collected and prepared data for multiple projects, ensuring data accuracy and reliability.

Statistical Analysis and Hypothesis Testing:

- I actively engaged in statistical analysis and hypothesis testing. This included conducting descriptive statistics, inferential statistics, and hypothesis testing to derive insights and support decision-making.

- Accomplishment: My statistical analyses and testing contributed to data-driven decisions that significantly impacted project outcomes, leading to actionable recommendations.

Data Visualization:

- I used data visualization techniques to present findings and insights effectively. This included creating charts, graphs, and visual representations of data to aid in the communication of results.

- Accomplishment: Through data visualization, I conveyed complex information in a clear and concise manner, facilitating better understanding among team members and stakeholders.

<u>Project Work:</u>

- I actively participated in data science projects that ranged from data analysis to predictive modeling. These projects required me to apply data science techniques, collaborate with the team, and meet project deadlines.

- <u>Accomplishment:</u> I played a significant role in the successful completion of data science projects, contributing to actionable insights and recommendations. These projects included customer behavior analysis, sentiment analysis, and predictive modeling.

<u>Collaboration and Teamwork:</u>

- Collaboration with the data science team was a crucial aspect of my role. I actively participated in team discussions, shared insights, and worked collectively on projects.
- <u>Accomplishment:</u> Effective collaboration resulted in the successful completion of data science projects, as well as a positive contribution to the team's collective knowledge.

<u>Challenges Overcome:</u>

- Challenges during the internship included handling and cleaning large and messy datasets, navigating through complex data structures, and managing project timelines.

- <u>Accomplishment:</u> These challenges were overcome through research, persistence, and collaborative efforts, resulting in successful project outcomes.

Overall, my internship at Skill vertex provided opportunities to apply theoretical knowledge to practical data science tasks, contributing to the organization's objectives and my personal and professional growth. The experience allowed me to tackle real-world data challenges and make data-driven decisions effectively.

V SKILLS LEARNED

Data Collection and Preprocessing:

I acquired practical skills in data collection and preprocessing, including data cleaning and transformation. This included techniques to ensure data quality and reliability, which are essential in real-world data science applications.

Statistical Analysis and Hypothesis Testing:

I developed proficiency in statistical analysis and hypothesis testing. This involved applying various statistical techniques to analyze data and evaluate hypotheses, which are foundational skills in data science.

Data Visualization:

I learned how to effectively visualize data by creating charts, graphs, and visual representations. This skill enables clear and concise communication of insights and findings to both technical and non-technical stakeholders.

Project Management:

Through project work, I gained experience in project management, including setting goals, managing timelines, and meeting deadlines. These skills are invaluable for efficient project execution.

Team Collaboration:

Collaboration with the data science team enhanced my teamwork and communication skills. Working collectively on projects improved my ability to contribute to a team's success.

Relevance to the Future:

These newly acquired skills are highly relevant to my future endeavors:

Professional Growth:

The skills in data collection, preprocessing, and statistical analysis are fundamental in the field of data science. They will serve as a solid foundation for my career and enable me to tackle complex data challenges effectively.

Data-Driven Decision-Making:

The ability to conduct statistical analysis and hypothesis testing is essential for making data-driven decisions in various industries. These skills will enable me to provide evidence-based recommendations and insights to support organizational goals.

Data Communication:

Data visualization skills are invaluable for presenting data-driven insights in a clear and compelling manner. This will be beneficial in communicating findings to both technical and non-technical audiences, making data more accessible and actionable.

Project Management and Teamwork:

Project management and teamwork skills will be useful in my future career. These skills facilitate effective collaboration with colleagues, enable efficient project execution, and contribute to overall project success.

Relation to Academic Background:

My academic background provided the theoretical foundation for the skills I acquired during this internship. The knowledge gained in statistics, data analysis, and data science methodologies at the academic level was applied practically in real-world scenarios.

This hands-on experience bridged the gap between theory and practice, enriching my academic foundation and making it more applicable to the professional realm.

Overall, the skills developed during this internship have contributed significantly to my personal and professional growth, equipping me to excel in data-driven professions and make meaningful contributions to organizations in the future.

VI TOOLS AND TECHNOLOGIES

Windows OS:

Windows is a **graphical operating system** developed by Microsoft. It allows users to view and store files, run the software, play games, watch videos, and provides a way to connect to the internet. It was released for both home computing and professional works.

YARN (Yet Another Resource Negotiator):

YARN is a resource management technology in Hadoop that facilitates job scheduling and cluster resource management, particularly for data processing and analysis.

Git Bash:

Git Bash is a command-line interface (CLI) for Git, a version control system. It provides a Unix-like terminal experience for Windows users, enabling them to use Git commands.

GitHub:

GitHub is a web-based platform for version control and collaboration using Git. It's widely used for software development, allowing multiple users to work on the same project.

Python:

Python is a versatile and high-level programming language known for its readability and simplicity. It's widely used in web development, data analysis, artificial intelligence, and more.

NumPy:

NumPy is a Python library for numerical and scientific computing. It provides support for arrays and matrices, along with mathematical functions to manipulate them.

Pandas:

Pandas is a Python library for data manipulation and analysis. It offers data structures like dataframes and tools for cleaning, transforming, and exploring data.

SciPy:

SciPy is an open-source Python library for scientific and technical computing. It builds on NumPy and provides additional functionality for optimization, integration, interpolation, and more.

Matplotlib:

Matplotlib is a Python library for creating 2D and 3D plots and visualizations. It's widely used in data analysis, allowing users to create graphs and charts.

Seaborn:

Seaborn is a data visualization library for Python. It's built on top of Matplotlib and provides a high-level interface for creating informative and attractive statistical graphics.

PySpark:

PySpark is the Python library for Apache Spark, an open-source big data processing framework. It allows data processing and analysis at scale using Python.

Tableau:

Tableau is a powerful data visualization tool that allows users to create interactive and shareable dashboards. It's used for exploring and understanding data, making it accessible to non-technical users.

These technologies and tools play significant roles in various domains, from data analysis and software development to system administration and big data processing. Each has its unique strengths and applications in the world of technology and data.

VII FUTURE GOALS

<u>Building a Career in Data Science:</u> The internship should serve as a stepping stone toward building a successful career in data science. Your future goal may involve securing a data science position or advancing to more senior roles within the field.

<u>Advanced Data Analysis:</u> The experience gained during the internship should pave the way for more advanced data analysis projects. Your goal may be to work on complex data problems, perhaps involving machine learning and artificial intelligence, to tackle real-world challenges.

<u>Contributing to Research and Innovation:</u> If research is a passion, you might aim to contribute to innovative data science research and publications, potentially making significant advancements in the field.

<u>Continuing Education:</u> Continuous learning is essential in the ever-evolving field of data science. Your goal may include further education, such as pursuing a master's or Ph.D. in a related field, or obtaining additional certifications in data science.

<u>Mentorship and Leadership:</u> As you gain experience, you may aspire to become a mentor for upcoming data scientists or take on leadership roles within data science teams or organizations.

<u>Data-Driven Decision-Making:</u> Applying your data science skills to influence strategic decisions within an organization can be a valuable goal. This might involve becoming a chief data officer or holding a similar leadership position.

<u>Community Involvement:</u> Giving back to the data science community by contributing to open-source projects, organizing data science events, or participating in initiatives to promote data literacy and ethical data practices.

VIII CONCLUSION

In retrospect, my Data Science Internship at Skill vertex from 5th April 2023 to 5th June 2023 has been a transformative and enlightening experience. This concluding section aims to summarize the main points of this report and discuss the overall value of this internship journey.

Throughout the internship, I set out with clear objectives, namely to acquire a comprehensive understanding of data science techniques, enhance data analytics skills, and apply statistical and hypothesis testing methods to real-world data scenarios. I am pleased to affirm that these objectives were not only met but exceeded.

The internship provided a platform for practical learning and hands-on application of data science concepts. I actively engaged in various tasks, from data collection and preprocessing to statistical analysis and data visualization. These experiences enriched my skill set and deepened my knowledge of data science fundamentals.

Collaboration with the data science team allowed me to contribute meaningfully to data-driven projects, aligning with Skill vertex's data strategy. The mentorship of experienced professionals was invaluable, guiding my journey and providing insights that extended beyond textbooks and academic theories.

The skills acquired during this internship, including data collection, statistical analysis, and data visualization, have equipped me for a future in data-driven professions. These skills will undoubtedly be instrumental in making informed, data-driven decisions and deriving actionable insights from complex datasets.

In conclusion, my internship at Skill vertex has been an enriching experience, bridging the gap between theory and practice. It has empowered me with a set of skills that I can carry forward in my academic and professional journey. I am grateful for the support and guidance I received, and I look forward to applying these skills in my future endeavors

IX ACKNOWLEDGMENTS

I would like to express my sincere gratitude to the individuals and groups who have played a pivotal role in making my Data Science Internship at Skill vertex an enriching and valuable experience.

First and foremost, I would like to extend my heartfelt appreciation to the entire Skill vertex team, particularly the data science professionals, mentors, and colleagues. Their guidance, support, and willingness to share their knowledge have been instrumental in my professional growth and the success of this internship.

I am also thankful to my academic institution for providing me with the foundational knowledge that enabled me to make the most of this internship opportunity. The theoretical background I gained during my studies served as a strong base upon which I could build practical skills.

My gratitude extends to the Skill vertex community and fellow interns who created a collaborative and inspiring environment. The shared insights, discussions, and teamwork have been invaluable in my journey.

I would like to thank my family and friends for their unwavering support, encouragement, and understanding throughout this internship. Their belief in my capabilities was a driving force in my accomplishments.

Lastly, I would like to thank all the resources, organizations, and platforms that provided valuable materials, tools, and documentation to aid my learning and project work during this internship.

This journey has been a collective effort, and I am deeply appreciative of every person and resource that contributed to my growth during this internship. I look forward to applying the knowledge and skills gained here in my future endeavors.

X APPENDIX

EMAIL FROM SKILL VERTEX

Skillvertex: Certificate of Program Completion Inbox ×

Info . <info@skillvertex.com>
to me ▾

Hi **SkillLearner!**

We congratulate you for successfully unlocking your certificate of completion for the program.
You can upload the certificate on Linkedin in order to reflect the authenticity of your acquired skills and add value to your profile.

Use **#skillvertex** to get an appearance with our partner companies to get the opportunities (paid internships/ jobs).

Kindly find the certificate in the mail attachment.

We wish you all the best in your future endeavors.

Kind Regards,

Team Skillvertex
Support@Skillvertex.com

+91 9606012806
+91 8095733636 | www.skillvertex.com

Working Days: Mon-Sat[11:00 AM-8:00 PM]

One attachment • Scanned by Gmail ⓘ

[PDF] data-7.pdf

TRAINING CERTIFICATE FROM SKILL VERTEX

SKILL VERTEX

CERTIFICATE
OF TRAINING COMPLETION

THIS CERTIFICATE IS PROUDLY PRESENTED TO

ARUNKUMAR.A

This is to certify that the above mentioned candidate has successfully completed his/her training in DATA SCIENCE from 5th April to 5th May 2023.

During this course he/she showed diligence, consistency, determination, active participation and innovation throughout their training period.

Mayank Gathole
(Academic Head)

Student UIN : **SV229199**

Certificate no. : **RA229699**

PROGRAM COMPLETION CERTIFICATE FROM

SKILL VERTEX

Internship Completion Certificate

This is to certify that *ARUNKUMAR.A* has successfully completed the Data Science course offered by SkillVertex beginning 5th May 2023 & ending 5th June 2023, which covers the fundamental concepts and applications of Data Science.

The course covered a wide range of topics, including data analysis, data visualization, statistical inference, machine learning, and other related fields. *ARUNKUMAR.A* has demonstrated a deep understanding of these concepts and has shown an ability to apply them to real-world problems.

During the course, *ARUNKUMAR.A* actively participated in all learning activities, including lectures, discussions, assignments, and projects. [He/She] consistently demonstrated a high level of engagement and critical thinking skills, which resulted in [his/her] outstanding performance throughout the course.

As such, *ARUNKUMAR.A* has met all the requirements of the Data Science course and is hereby awarded this certificate of completion.

We wish *ARUNKUMAR.A* all the best in [his/her] future endeavors and hope that the knowledge gained from this course will enable [him/her] to excel in the field of Data Science.

Issued on: 5th June 2023

Dinesh Singh
Head of Human Resource
SkillVertex