BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI
4th SEMESTER 2017-19 Batch
SE LOWES ZG629T DISSERTATION
<u>Dissertation Outline</u>

BITS ID No. <u>**2019AT15015**</u>   **Name of Student:** <u>**Aarti Pramod**</u>

Name of Supervisor: <u>**Siva Prasad Geetha Nagarajan**</u>

Designation of Supervisor: <u>Senior Technical Lead - Data Engineering & ML</u>

Qualification and Experience: <u>Masters in DataScience , 11 years</u>

E- mail ID of Supervisor: <u>sivaprasad.nagarajan@atmecs.com</u>

Topic of Dissertation: <u>Exploratory Data Analysis (EDA) using python on Women's health and fitness - A case study.</u>

<u>Name of First Examiner:</u> <u>Prof. Gururaj H S</u>

Designation of First Examiner:_____

Qualification and Experience: _____

E- mail ID of First Examiner: <u>gururaja@wilp.bits-pilani.ac.in</u>

<u>Name of Second Examiner:</u> <u>Prof. Mohammad Saleem Bagewadi</u>
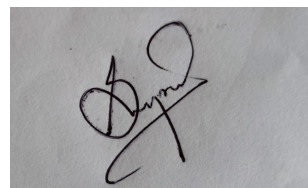
Designation of Second Examiner:_____

Qualification and Experience:_____

E- mail ID of Second Examiner: <u>mj.bagewadi@pilani.bits-pilani.ac.in</u>

(Signature of Student)   (Signature of Supervisor)

Date: 2023-12-15   Date: 2023-12-15

Index :

1. Dissertation Topic :

<u>Exploratory Data Analysis (EDA) using python on Women's health and fitness - A case study.</u>

2. Dissertation Title

Unveiling Health and Fitness Insights through Exploratory Data Analysis: A Python-Powered Approach

3. Objectives

1. To conduct an in-depth exploration of health and fitness data using Python programming language.
2. To identify patterns, trends, and relationships within the data that can provide insights into health and fitness behaviours.
3. To develop visualisations and data summaries that effectively communicate the findings of the analysis using streamline library.

4. Scope of work

1. **Data Collection:** Gather a comprehensive dataset of health and fitness data from various sources, such as wearable devices, fitness apps, and healthcare records.
2. **Data Preprocessing:** Clean and prepare the data for analysis, handling missing values, data inconsistencies, and data transformations.
3. **Exploratory Data Analysis (EDA):** Employ Python libraries and techniques to perform EDA, including descriptive statistics, data visualisation, and correlation analysis.
4. **Data Visualisation:** Create informative and engaging visualisations using Python libraries like Matplotlib, Seaborn, plotly and Bokeh to represent the data insights.
5. **Data Summaries:** Summarise the key findings and insights obtained from the EDA, highlighting patterns, trends, and relationships within the data.

5. Background of previous work done in the chosen area

Exploratory Data Analysis (EDA) is a commonly used technique to find hidden patterns, trends, and relationships within complicated datasets. Examples of these domains include healthcare and fitness. EDA approaches have been employed by researchers to obtain insights into the elements that impact health outcomes, pinpoint risk factors associated with chronic illnesses, and formulate customised health interventions.

As a dedicated data engineer of my team I have had the privilege to use python and analyse a wide range of data and showcase the analysis.

6. Methodology (ways and means to be adopted in achieving the objective/ aim)

1. **Data Collection:** Utilize various data sources, including wearable devices, fitness apps, healthcare records, and publicly available datasets, to gather a comprehensive collection of health and fitness data.
2. **Data Preprocessing:** Employ Python libraries like pandas and NumPy to clean and prepare the data for analysis, handling missing values, data inconsistencies, and data transformations.
3. **Exploratory Data Analysis (EDA):** Conduct EDA using Python libraries like pandas profiling, pandas-plot, data visualization, and correlation analysis.
4. **Data Visualization:** Create informative and engaging visualizations using Python libraries like Matplotlib, Seaborn, and streamlit to represent the data insights.
5. **Data Summaries:** Summarize the key findings and insights obtained from the EDA, highlighting patterns, trends, and relationships within the data.

Detailed Plan of Work (according to the semester calendar)

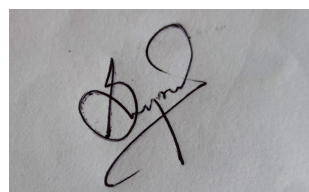| Serial Number of Task	**Tasks or subtasks to be done** (be precise and specific)	Expected date or week of completion	Specific Deliverable in terms of the project
Project Inception and Planning	Requirements gathering and discussions with supervisor including : 1. Data source 2. Design pattern 3. Platform to be used 4. Libraries to be used 5. Prototype of the application (poc)	Week 1-4 (Dec 4 - Dec 30)	Prototype of the project.
Development	Developing the full functionalities of the application.	Week 4-6 (Dec 31 - Jan 21)	Completed features and functionalities
Supervisor Feedback and improvisations	Code review with supervisor and revisions on the development.	Week 9-11 (Jan 22 - Feb 11)	Developed project ready to be deployed or hosted

Bug fixes; and Final Report Preparation and submission	1. Feedback and final report from supervisor 2. Examiner's feedback	Week 15-16 onwards (Feb 26 - March 14)	Final Report Preparation and submission.

Supervisor's Rating of the Technical Quality of this Dissertation Outline

EXCELLENT / GOOD / FAIR/ POOR (Please specify):   GOOD

Supervisor's suggestions and remarks about the outline (if applicable).

Date: 2023-12-15

(Signature of Supervisor)