Fitness Dataset Approach Report

Personal Fitness Dataset

Stephen Doan – Nguyen  
01/26/2024

Table of Contents

[Introduction 3](#_Toc163579931)

[Phase 1: Planning & Design 3](#_Toc163579932)

[Phase 2: Data Collection and Entry 4](#_Toc163579933)

[Phase 3: Data Cleansing 5](#_Toc163579934)

[Phase 4: Data Analysis 5](#_Toc163579935)

[Phase 5: Data Modeling and Visualization 6](#_Toc163579936)

[Phase 6: Analysis Report 7](#_Toc163579937)

[End Phase: Database System Project 8](#_Toc163579938)

[Conclusion 8](#_Toc163579939)

# Introduction

The Fitness Dataset approach outlines a meticulous design framework that presents the objectives, Key Performance Indicators (KPIs), and methodologies essentials for devising and executing creating a robust dataset. The comprehensive strategies delineate each phase’s objectives, setting clear benchmarks for the successful completion of each of the stages. Each phase delves deep into the approach and articulates the expectations and goals, ensuring a structured and linear progression toward achieving the overarching aim

By establishing a well-defined outline, the approach methodically demonstrates the phases of crafting a detailed dataset. The primary goal is to create a narrative and insightful methodology for the Fitness Dataset initiative. Moreover, the strategy emphasizes the importance of iterative refinement and adaptation during the phases, allowing the dataset to evolve during the process of the new findings and emerge new fitness adjustments

In essence, the Fitness Dataset Approach goal is to define the outline and framework so the Fitness Dataset transcends more than mere data collection but embodies the holistic journey towards understanding and optimizing fitness endeavours. It stands as a testament to the power of data-driven insight in forging a path of improving oneself each day by using curiosity about data.

# Phase 1: Planning & Design

Planning & Design sets the foundation for the Fitness Dataset Approach, defining the preparation and design for the structure. The goal of creating the Fitness Dataset has many functionalities from the simplest as tracking weight instead of memorizing to the more complex functionalities of determining how the dataset can improve one’s fitness regime. The planning & design main goal is to create a flexible long-term design that has many different functionalities as I’ve recognized that one’s fitness goal can change over time and thus can cause complexity if the structure is a strict design. Thus the periods of time will be split into blocks, approximately 2 months. This causes enough time to determine if the fitness regime is effective or not, or needs to be change. Creating blocks helps understand the reason for change for phase 6: Analysis as well it becomes more effective when cleansing the data for inconsistencies. Key data points in the structure is defined by the different muscle groups, this ultimately separates and dwells more in-depth in the data points. The Data Structure revolves completely around ones personal fitness schedule, specifically following a push-pull-leg split. The structure also includes other variables to create a more accurate representation and more insightful analysis for future processes and phases. The other key defining variables are intensity, weight lifted, date, reps and sets, and other small contributing points

# Phase 2: Data Collection and Entry

Data Collection and Entry’s goal is to efficiently and effectively populate the dataset using primary and relevant data. This involves meticulously approaching the data entry ensuring that each piece of data aligns perfectly with the structure and the other predefined data points, preventing any inconsistencies in the structure. The time phase is approximately 10-12 months ensuring a large enough data set to provide accurate insights and findings of ones workout regime. Simultaneously while entering relevant data, rigorous attention is needed to maintain the format and data validation from human errors, reducing potential errors for inconsistencies for further process in the phases. Moreover, a key focus of this phase is to optimize the process through automation wherever possible instead of manually entering data for the possibility of human error. This is possible by leveraging tools and apps that can automate the entry of data in the form of fitness apps or such.

Ultimately, Phase 2 is about establishing a robust, reliable, and efficient workflow to implement the data into the dataset. Achieving these goals ensures the foundation for future processes

# Phase 3: Data Cleansing

Phase 3: Data Cleansing is a pivotal stage in the Fitness Dataset, this is where the integrity and usability of the collected data are refined and assured for future analysis and processes within the stage. The goal of the phase is to enhance the dataset’s quality, ensuring it is pristine, reliable, and ready for analysis in phase 4: Analysis. The process involves a meticulous process of examination of the data to identify and rectify any null, duplicates, errors, irrelevant entries, and inconsistencies that may have occurred during phase 2: Data Collection and Entry. Any discrepancies left undressed can cause major implications during the analysis and could skew analysis leading to inaccurate conclusions and findings. The approach starts with the identification of problematic data with the combination of automated tools as well as manual review. Scanning for empty cells (nulls), verifying data against known standards, identifying duplicate records, and filtering out any data points not aligning with the objectives or structure of the dataset. Moreover, ensuring consistency across the dataset is crucial, especially in terms of naming conventions and units, which can facilitate a smooth analysis in the later phase.

# Phase 4: Data Analysis

Phase 4: Data Analysis is the analytical core of the Fitness Dataset, where the gathered and cleansed data is transformed into actionable in-depth insights. The goal of this phase is to conduct a comprehensive analysis that quantifies the current state of fitness metrics but also uncovers deeper patterns and relationships within the data, aligning with the KPIs and objectives. The process involves a two-tiered approach: starting with basic Excel calculations to compute the averages, sum, and other straightforward metrics that provide an initial overview of the dataset. The basic analysis offers valuable insight and a foundation for the second tier, setting the stage for more complex investigations.

The second tier approach follows more of an advance and sophisticated analysis aiming to drill down the specifics of the dataset, utilizing more complex tools such as PivotTables, Visual Basic Automation, and Toolpak. Pivot tables to facilitate the dynamic summarization of the exploration of the data, providing a detailed dimension of categories and variables. ToolPak gives the ability for statistical tests, regression analysis, and complex calculations that are essential for identifying correlations, relationships, and patterns within the dataset. The advanced analysis is targeted towards the objective and KPIs.

Furthermore, the phases place great importance on the correlation and relationship analysis on the understanding of how different data points can influence on another as well as applying and observing any trends or anomalies. Phase 4 transforms raw data into a comprehensive narrative of the fitness progress and its challenges, forming a path for decisions.Top of Form

# Phase 5: Data Modeling and Visualization

Phase 5: Data Modeling and Visualization is the transition of the analytical work of Phase 4 into compelling visual narratives graphically. The primary goal is to make complex insights from the data analysis phase accessible and understanding, to find even deeper findings. This can be achieved through charts, graphs, and dashboards that represent the data and its findings graphically. By creating visual representations the approach to seeking key trends, patterns, and anomalies uncovered during the analysis make a easier comprehension of the result

A primary focus of Phase 5 is developing detailed dashboards that summarize the key metrics, KPIs, and goals mentioned in Phase 1. The dashboards are designed to be interactive and informational to allow users to drill down closer to the specific aspects of the data, evaluating the intricate portions of the data.

The approach incorporates advanced visual tools like Power of BI and Tableau, as these platforms offer a wide array of visualization options and flexibility to handle large datasets, making them suitable and ideal for creating dynamic and insightful dashboards. The use of Power BI and Tableau extends more than pure aesthetic representation, but enables the integration of data from various sources, ensuring a seamless and unified analytical experience.

Phase 5 is critical for bridging the gap between raw data and actionable insightful knowledge, enabling informed decisions that can improve ones fitness regime and overall health outcomes

# Phase 6: Analysis Report

Phase 6: Analysis Report is the final phase of the project, dedicated to consolidating and communicating the insights that are derived from the preceding phases. The core objective is to create a coherent, comprehensive report that not only documents the findings on relationships, correlations, and patterns among the various data points but also provides a summary of identified and actionable recommendations for my personal fitness regime. This phase also involves the meticulous process of the analytical journey, highlighting how the data was transformed into meaningful insight through the process outlined in the earlier phases. The approach of the Analysis report is for it to be both narrative and analytical, aiming to tell the my story through the process of data. How initial goals and KPIs led to specific discoveries about fitness trends, behaviours, and the result. It underscores the significance of the relationship and correlation, interpreting these findings within the context of my fitness regime. The report also includes a section dedicated to summarizing these findings and decisions on how I can adjust my fitness regime for improved results, based on the gathered insights from the data.

By encapsulating the journey and its outcome, the Analysis Report stands for a testament to the value of data-driven decision-making in personal fitness. It not only offers guidance and recommendations but also sets a foundation for ongoing analysis and refinement, fostering the culture to improve one’s self.

# End Phase: Database System Project

The End Phase is actually the beginning of a new phase of a project. It is to represent the dataset in a database approach transitioning from a spreadsheet-based analysis to a more structured, scalable, and robust framework using a Relational Database Management System utilizing SQL queries. This phase enhances the dataset’s integrity, accessibility, and analytical capabilities. This end phase is the beginning of a new phase to open new routes of data-driven decisions and to ultimately fulfill my curiosity of data in a new environment.

# Conclusion

The purpose of the report is to create a seamless transition between the phases and the workflow process ensuring a successful completion by defining the framework and outline of the approach. By leveraging the approach and phases, we can clearly articulate and define clearer goals of the project leading to a more in-depth analysis of the goals. The dataset not only has the function to articulate analysis for decision but can tell a story and this is the story of me improving myself, one step at a time!