

**IFN701 Project 1 – Project Plan**

**Symptoms Of Lower Back Pain: A Data Analysis Project**

Davinder Kaur

N9617761

**Supervisor:**

Dr. Guido Zuccon

**1.0: Introduction**

***Background:*** inct fields

Lower back pain data analysis new practice to understand the symptoms of this problem

Here defines the value of data analysis

Introduction of this project dataset:

310 observations

Contains 12 attributes

* 1. **Significance:**
  2. **Stakeholders:**

|  |  |
| --- | --- |
| **Stakeholder** | **Participation** |
| Student | Responsible for project development |
| Supervisor | Give Suggestions and direct the student |
| Project Coordinator | Determine project progress |

* 1. **Deliverables:**

The main motive of this project is to do data analysis by available practices using R. The tangible outcomes of this data analysis project are illustrated below:

* **Data analysis report:** This data analysis report holds the delated information about statistical summary, statistical testing (Null hypothesis / alternative hypothesis). Apart from this, it also reveals the relationship between various symptoms. It reports the trends may observed in distinct symptoms/ visualizations.
* **R markdown:** The documents created in R Studio are Known as R markdown. These R markdowns will provide all the source code that is used for data analysis. Also, markdown would contain data visualization, prediction model, and confusion matrix.
  1. **Project Scope:**

The following table of project scope will provide the clear picture of what would be included and excluded in this project.

|  |  |
| --- | --- |
| In Scope | Out of Scope |
| * Literature Review for the sake of problem comprehension * Analyse the dataset using R programming * Creating R markdown * Evaluate the relationship present between various attributes/Symptoms * Building prediction model * Communication via data analysis report | * Recommendations to solve this problem would be out of scope * Building new dataset for data analysis * Any remedial suggestion to alleviate the problem |

**2.0: Project methodology:**

These are the seven steps are necessary that should be implemented to achieve the target deliverables of this project.

**Step 1:** Firstly, the Literature review is necessary for the comprehension of the problem space. Though, this issue is broad, narrowing the scope of project may helpful by extending knowledge from other sources. Also, the literature review also aids in thought process to derive new ideas.

**Step 2:** Most commonly, data analysis projects start with defining the ideal Dataset, which act as a fundamental step for extracting the information in order to solve the problem or to gain insights.

**Step 3:** For this project, the dataset is downloaded from “Kaggle” website. The Kaggle is well known for its data analysis competitions, were business people and researchers upload their data. The datasets using for this project is in csv (comma separated file) format.

**Step 4:** Data cleaning is a vital step of data analysis workflow. It is very essential to get accurate and reliable results. In this phase, problems need to solve such as dropping unnecessary data, evaluating missing/ incomplete data/ noisy data or any kind of inconsistencies.

**Step 5:** To deal with, dirty data some data cleaning steps will be required. However, data exploration is equally important. To cite an example, if the dataset holds null values then removing those data point can be beneficial because null values cannot help in describing pattern. Few other approaches that

**Step 6:** Creating hypothesis and appropriate visualization the data using R will be done. This, hypothesis

**Step 7 and 8:** For better understanding, interpretation of result will be elaborated in the markdown as well as in data analysis report.

**3.0: Project Management Approach:**

In order to manage this project, Dynamic Systems Development Method (DSDM) will be applied. Basically, it is a method from agile framework that not only supports on time delivery of project but also easy to manage. The iterative, continuous incremental delivery, and collaboration are the core principles of DSDM project management. Moreover, DSDM increases the opportunities to receive feedback from the Supervisor. As a result, it effectively caters requirements of on project that are indispensable for the quality and control management. The DSDM Agile Project Framework (2014 Onwards)(n.d.). In Addition, the principles are considered as the foundation of this methodology that makes it more suitable for this project. Those are:

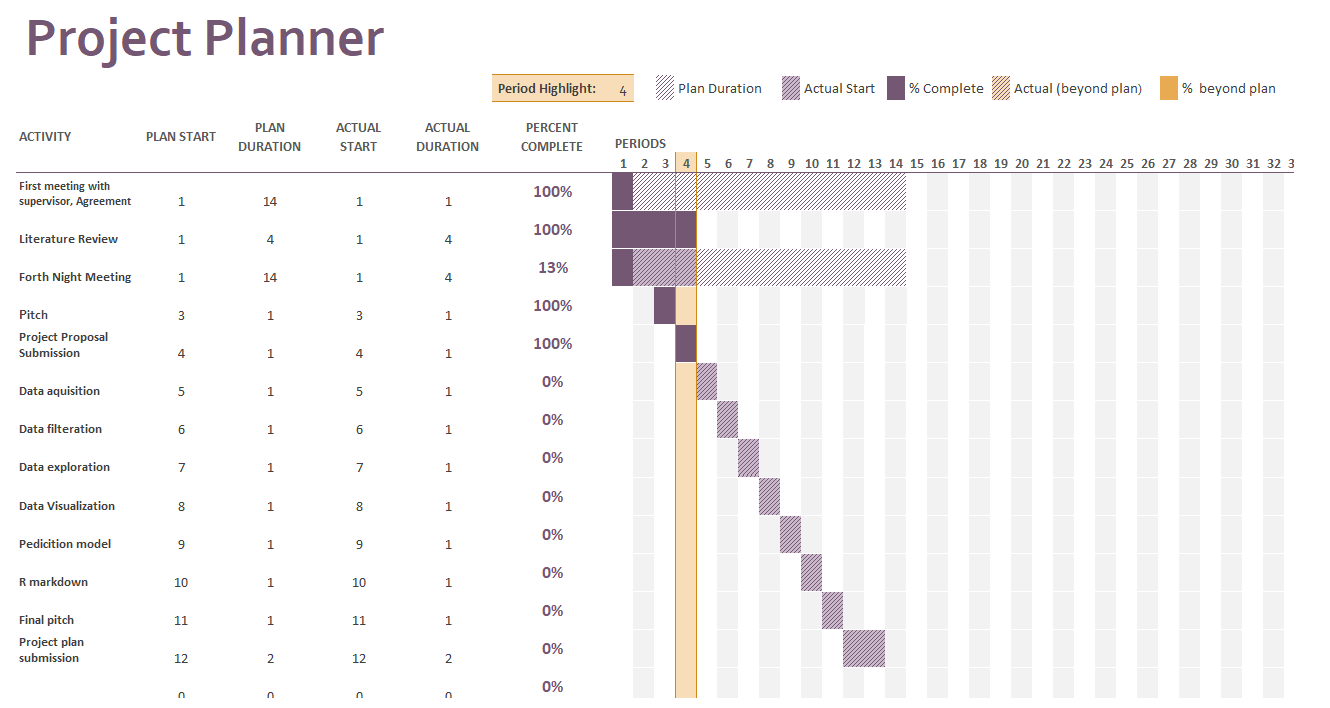
* Focus on project needs
* On time delivery
* Build incrementally
* Communicate constantly
* Quality assurance
* Iterative development
* Transparency in work

**3.1: Project Timeline:**

|  |  |
| --- | --- |
| Timeline | 13 weeks |
| Number of Increments | 3 (1 month each) |
| Number of Timeboxes | 6 (2 weeks each) |
| Time-period of each time box | 5 days |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Kick off | Close out | Planned activities | Tasks/Outcome | |
| *Increment 1* | | | | |
| Timebox 1 | | | | |
| Week 1 | | * Arrange meeting with supervisor and project discussion * Get familiar with GitHub | | * Sourced project * Installation of Rstudio * Download the Dataset * Join Slack group |
| Week 2 | | * Sign off the project agreement * Analysis for problem space | | * Agreement Submission * Repository on GitHub * Brief Literature review |
| Timebox 2 | | | | |  | |
| Week 3 | | Prepare project presentation | | * Project proposal pitch |
| Week 3 | Week 4 | Work on Project Proposal | | * Proposal submission |
| *Increment 2* | | | | |
| Timebox 3 | | | | |
| Week 5 | | Data Analysis | | * Data acquisition |
| Week 6 | | Data Analysis | | * Data cleaning |
| Timebox 4, | | | | |
| Week 7 | | Data Analysis | | * Data exploration |
| Week 8 | | Data visualization | | * Data visualization |
| *Increment 3* | | | | |
| Timebox 5 | | | | |
| Week 9 | | * Prediction modelling * Registration for week 12 presentation | | * Data modelling and statistical prediction * upload slides on GitHub |
|  | | **Mid Semester break** | |  |
| Week 10 | | * prepare project plan presentation * Knit/Create R markdown and upload on GitHub | | * refinements in R markdown |
| Timebox 6 | | | | |
| Week 11 | | * Documentation of data analysis * Submit the draft on GitHub for feedback | | * Data Analysis report * Project plan presentation |
| Week 12 | week 13 | Work on final report | | * Final report submission |

**3.2: Gantt Chart:**

****

**3.4: Prioritization:**

**3.5: Communication Plan:**

The communication plan presents information that is vital for people, who are engaged in this project. It elaborates in terms of classification, purpose, methodology, frequency, deliverables, and participants. The details of communication plan are given below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classification | Purpose | Method | frequency | Deliverables | Participants |
| Weekly meeting | Thorough appraisal of current timebox and planning for next timebox | Face to face | forth night or as needed | * Log sheet | * Academic supervisor * Student |
| Collaborative writing | To work closely with the Supervisor and take feedback.  Working in the assigned Lab at University | Face to face / GitHub/ Slack | Weekly | * Partial Solution in chunks | * Academic supervisor * Student |
| Notifications | Information about mandatory ToDos, clarification of doubts. Also, receive help from supervisor | Slack | Casual | * Chronicle of communication via slack and planned series of action | * Academic supervisor * Student |
| Communicate through E-mail | To deal with unexpected problem or to inform about formal meeting | Email | Frequent, rely on problem | * Records of continuous communication | * Academic supervisor * Student |
| Review session | Report errors and get feedback for  necessary  improvements | Face to face | After termination of each increment | * Log sheet | * Academic supervisor * Student |

**5.0: Risk Assessment:**

This section comprises all the potential risks that may occur during the life cycle of the project, which can influence the delivery of the project. The key risks are classified according to ranking of (High, medium and low). Basically, assessment of these risks not only stimulates the on-time delivery of project but also hike the overall workflow efficiencies. The Mitigation strategies aids in how to addresses the certain risk. The main risks are shown below:

|  |  |  |
| --- | --- | --- |
| **Risk Ranking (High, Medium, Low)** | **Risk Description and Impact** | **Mitigation Strategy and/or Contingency Plan** |
| **Medium** | Unavailability of Supervisor may lead to insufficient guidance | Set-up alternatives for communication likewise Slack or video conferencing. |
| **Medium** | Deliverables are not what exactly academic supervisor wants. Thus, it increases the likelihood to deliver the wrong solution. | Deliver the partial solution in small manageable parts. |
| **High** | As communication act an essential element in project management. Consequently, shortage of communicating plans can result in arguments. | Effective communication plan like weekly meeting, minutes meeting etc. |
| **Low** | Project goes beyond the set time limit (13 weeks) can affect the on-time completion of project | Project plan must be followed strictly and occurrence of issue needs to be resolved as soon as possible |

**References:**

The DSDM Agile Project Framework (2014 Onwards)(n.d.). Agile Business Consortium Retrieved from <https://www.agilebusiness.org/content/principles>

**Appendix:**

Feedback from oral presentation: