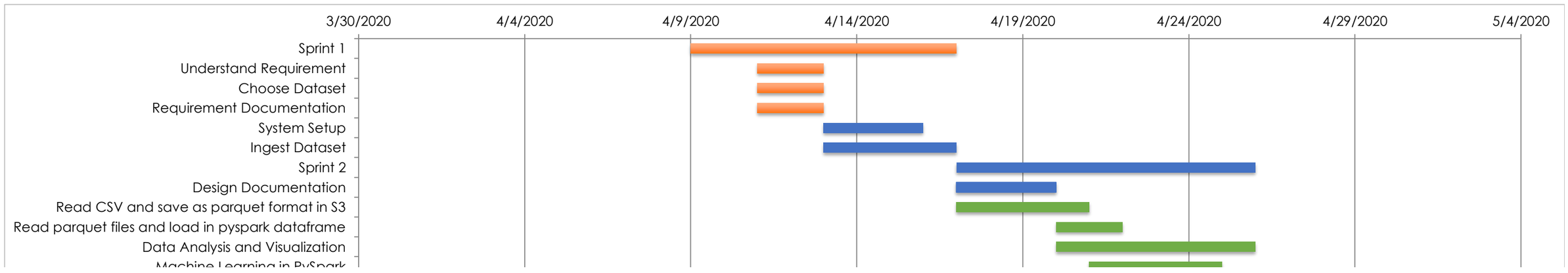
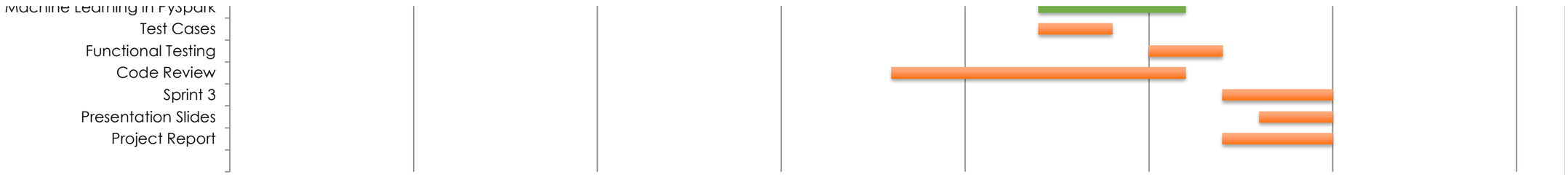
Gantt Chart

PROJECT NAME	PROJECT TEAM	START DATE	END DATE	OVERALL PROGRESS
Unstructured and Distributed Data Modelling Project	Team Blanca Peak	4/9/2020	4/28/2020	70%

AT RISK	TASK NAME	FEATURE TYPE	STORY POINTS	START	FINISH	DURATION (DAYS)	STATUS	COMMENTS
☐	**Sprint 1**			4/9/2020	4/16/2020	8	Completed	
☐	Understand Requirement	Requirement Analysis	1	4/11/2020	4/12/2020	2	Completed	
☐	Choose Dataset	Requirement Analysis	1	4/11/2020	4/12/2020	2	Completed	
☐	Requirement Documentation	Requirement Analysis	1	4/11/2020	4/12/2020	2	Completed	
☐	System Setup	Development	1	4/13/2020	4/15/2020	3	Completed	
☐	Ingest Dataset	Development	3	4/13/2020	4/16/20	4	Completed	
☐	**Sprint 2**			4/17/2020	4/25/2020	9	Completed	
☐	Design Documentation	Application Design	2	4/17/2020	4/19/2020	3	Completed	
☐	Read CSV and save as parquet format in S3	Development	3	4/17/2020	4/20/2020	4	Completed	
☐	Read parquet files and load in pyspark dataframe	Development	2	4/20/2020	4/21/2020	2	Completed	
☐	Data Analysis and Visualization	Development	3	4/20/2020	4/25/2020	6	Completed	
☐	Machine Learning in PySpark	Development	3	4/21/2020	4/25/2020	4	Completed	
☐	Test Cases	Testing	2	4/21/2020	4/22/2020	2	Completed	
☐	Functional Testing	Testing	1	4/24/2020	4/25/2020	2	Completed	
☐	Code Review	Review	1	4/17/2020	4/25/2020	8	In Progress	
☐	**Sprint 3**			4/26/2020	4/28/2020	3	TBD	
☐	Presentation Slides	Documentation	2	4/27/2020	4/28/2020	2		
☐	Project Report	Documentation	2	4/26/2020	4/28/2020	3		
☐								



Gantt Chart


Machine Learning in Pyspark
Test Cases
Functional Testing
Code Review
Sprint 3
Presentation Slides
Project Report