



Operation 3SL

Slow Speed Saves Lives

Phase 4

System Implementation

CIS 9590

Section QMWA

Suhaib Alsahybi

Sage Ashique

John Ortiz

Paridhi Sharma

Eren Yuceer

suhaib.alsahybi@baruchmail.cuny.edu

abdur.ashique@baruchmail.cuny.edu

john.ortiz1@baruchmail.cuny.edu

paridhi.sharma@baruchmail.cuny.edu

ismail.yuceer@baruchmail.cuny.edu

Phase 4. System Implementation

Executive Summary	2
Project Summary	3
Hardware/Software	8
Testing/Training	9
Support Plan	11
Implementation Plan	13
Variance from Original Specifications	13
Proposed Modifications	14
Decision Log	15
Appraisal Sheet	16

1. Executive Summary

The launch of NYC's Vision Zero, an initiative created to reduce traffic accidents and stop traffic fatalities, helped decrease the number of city wide accidents and traffic related deaths.

However, over the last 3 years, a decrease in the efficiency of the program required new measures to be taken by the NYPD.

Studies have shown that once drivers detect cameras in a certain area, they will consciously slow down in order to avoid a ticket. However this does not deter speeding in other areas. As such, as accidents decrease around intersections where camera locations are no longer a surprise, accidents continue to increase in other areas.

Project O3SL (Slow Speed Saves Lives) was established in order to add a new mobile layer of surveillance in the most accident prone areas of New York City. The prime objective of O3SL has been to create an interactive dashboard (with a recommendation engine) as well as a report generation feature. The features within the dashboard should allow NYPD obtain a measure of the traffic dangers in different areas. With this added insight, NYPD can more effectively monitor dangerous intersections and streets, and achieve the goal of a 10% yearly reduction in fatalities through 2021.

The O3SL project, in its final phase, has been created with strict user requirements in mind. Utilizing a mix of data collection from NYPD's traffic accident database, an end to end analytics platform for data analysis, and other features such as street level mapping, top-10 geo-map recommendations, and real time report generation, O3SL is a powerful tool for the NYPD to help curb citywide accidents and fatalities.

2. Project Summary

- a. **Problem Statement:** The existing system in NYPD and DOT is mostly manual. Analyst within these organizations used complaints, calls and police reports reports to create data reports. Accident reports, speeding ticket information, Citizen calls and complaints are gathered and entered into NYPD database, it is accessible to the NYPD as a whole. The current usage of the data have the drawbacks listed below:

I- The reports that are wanted by the NYPD always have to be analyzed every time a report is requested. It was also a huge time sink since every report takes time to collect data, analyze it, and process it to make it presentable to the person who requested the data.

II- The data entered into the system is not always clean. The current data through 2018 has missing values in zip codes, addresses or time value. Cleaning that kind of corrupt data manually creates financial and time loss.

III- There is no top-down view of statistics at a glance. There are some general reports that generated such as accidents in 2018 or in a specific month in a specific year. Yet the requested statistics only contains the data through the date that report is requested. It is not live.

- b. **Solution Statement:**

O3SL offers a reporting tool that always gives the live results and prevents the problems addressed above on NYPD side.

I - With O3SL, it takes just a few minutes for NYPD to request the visualized generic reports and download them. The visualized data in O3SL dashboard can be reached by the user anytime via a web browser. There is no wait time to see the data.

II - The data entered into the system by NYPD manually is cleaned by a Python code. There is a algorithm behind of the O3SL which cleans the corrupt that containing missing values. Each data is cleaned automatically and then goes to dashboard. It eliminates the manual process in cleaning.

III - O3SL is developed to propose the reports that is mostly wanted and used by NYPD such as street level mapping, top-10 geo-map. Since the O3SL dashboard is fed by live data, the reports can be requested any time and it reflects the updates.

c. **Boundaries:**

One of the biggest boundaries in this project is that O3SL is working with the government related organizations such as NYPD and DOT. It is hard to change the working style of governmental organizations since they perform the same process over many years. At Operation 3SL we created the software tool carefully not to create any change in the process of how NYPD work and add a value into their organization without changing how they operates. The most important boundaries within the development of O3SL is listed at below:

- I. Complying with NYPD's existing processes: The tool has to comply with how NYPD operates. There is a certain way that NYPD collects data and enters them into its database. We do not require a change to that or request to add a column to the data sheet since it would create a big change. Therefore we need to create a tool which does not change anything in NYPD operation but adds value compare to their needs and necessities.
- II. Data: We only have access to open data of NYPD. Our data set is limited to what is shared as public. It has pros and cons. Cons are that the data which is not shared with public cannot be added to the Dashboard and we do not know how valuable they might be. The pros are that there is already a huge data available of the accidents and fatalities of the data is open to public and it is not a huge drawback. The sample size is large enough to create analysis from.

d. **Survey & Interview:**

In the first step of the project, we tested what kind of product the stakeholders want to see at the end since the product that would be given should be comply with the needs of the current problem. Therefore, before designing the actual product, the survey and interview was prepared to share with stakeholders to see what is wanted in the O3SL dashboard.

Survey:

1. What do you want us to consider while geographically mapping crash data?

☐ Injuries

☐ Fatalities

☐ Both

2. The mapping of the crash data should include accidents involving

☐ Pedestrians

☐ Bicycles

☐ Motor vehicles

☐ All

3. How do you want to visualize the crash data locations on the map? Select one or more

☐ Based on zip code

☐ Based on Borough

☐ Based on Precinct

The intention in the above survey was to understand the kind of data the stakeholders want to filter and can take report based on Zip Code, Borough, and Injuries/Fatalities.

After the survey is done, the interview at below is made within the organization to understand the vision of the project, success criterias, what would be different from the existing system and who would be the users. The aim of the interview was to make sure that everyone is in same page with the project's goal and vision.

Sample Interview Questions	
Project Vision	What is your vision for this project? What defines success for this project? What are the potential pitfalls?
Users	Who are the primary users of this tool?
Existing Tools and Methodologies	Is there any existing tool in use? What are their relative strengths/weaknesses? How is this offering different?

e. Data Gathering & Data Cleaning:

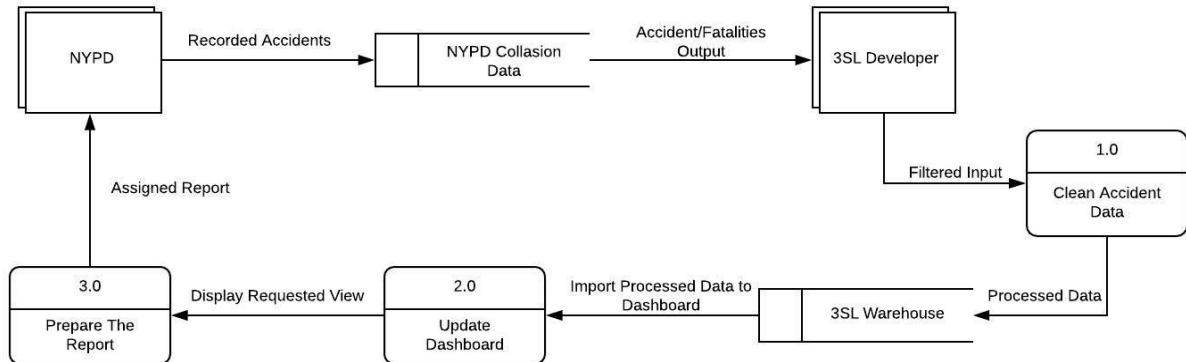
The result coming from the Survey is used in collection of data and how to use the data. NYPD-Motor-Vehicle-Collisions data was the main data source and there were twenty three of column which are Unique Key, Date, Time, Borough, Contributing Factor Vehicle 1, Contributing Factor Vehicle 2, Contributing Factor Vehicle 3, Contributing Factor Vehicle 4, Contributing Factor Vehicle 5, Cross Street Name, On Street Name, Off Street Name, Zip Code, Latitude, Longitude, Number of Cyclist Killed, Number of Motorist Killed, Number of Pedestrians Killed, Number of Person Killed, Number of Cyclist Injured, Number of Motorist Injured, Number of Pedestrians Injured, Number of Persons Injured.

The NYPD Collision Data is cleaned according to if the data has any missing values, combining the whole values under fatalities/injuries, these 23 number of field is filtered, the ones that would not be used in dashboard is eliminated by coding and reduced to 11 number of fields. As part of our data cleaning process, we removed attributes and record types which did not add value to the requirements of this dashboard.. The data fields chosen to store in data O3SL warehouse is listed at below:

Variable	Description	Datatype
date	Date of incident	date
time	Time of incident	time
borough	Borough of incident	catergorical
zip_code	Zip code of incident	catergorical
latitude	latitude coordinate of incident	float
longitude	longitude coordinate of incident	float
on_street_name	Street name within the legal driving area	string (object)
cross_street_name	Intersection crossing	string (object)
off_street_name	Street name Off of legal driving area	string (object)
fatalities	total # of fatalities that took place	integer
injuries	total # of injuries that took place	integer

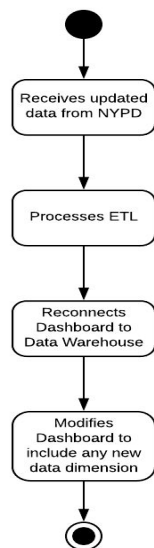
f. Design:

How all these data will flow from NYPD to O3SI and flow back to NYPD is designed and major data flow processed and determined below:

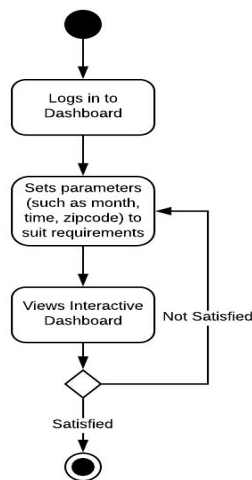


According to the Data flow diagram, there are two major actors in the system : O3SL developer and NYPD/DOT with different access rights to the dashboard. The differentiation between the users is shown at below Activity Diagrams:

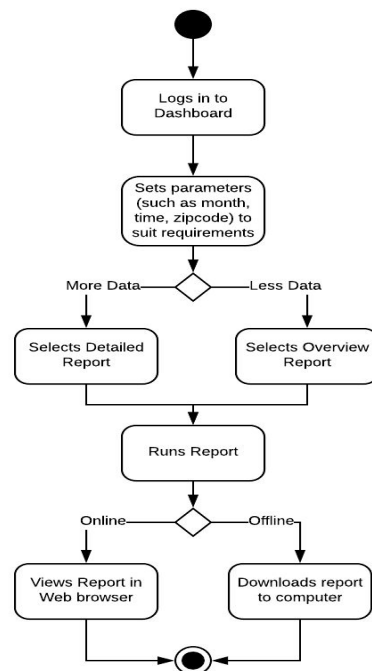
1. Tableau Developer Modifying Dashboard



2. NYPD or DOT Viewing Dashboard



3. NYPD or DOT Running Report



3. Hardware/Software

- a. List the hardware you are recommending for this project with costs
- The user and admins need to have access to the computer that supports given specification for Tableau Desktop.

Minimum System Requirements	
Windows	Mac
Intel Core i3 or AMD Ryzen 3 (Dual Core) 4GB or larger 2GB HDD free or larger*	Intel Core i3 (Dual Core) 4GB or larger 2GB HDD free or larger*
Recommended Requirements	
Windows	Mac
Intel Core i7 or AMD Ryzen 7 (Quad Core) 16GB or larger 2GB SSD free or larger*	Intel Core i7 (Quad Core) 16GB or larger 2GB SSD free or larger*
High Performance Requirements	
Windows Intel Core i7 or AMD Ryzen 7 (16 Core) 32GB or larger 2GB SSD free or larger*	

- Google Cloud platform to host data in warehouse after ETL process.

Storage Size: 1000GB

Data Retrieval Size: 1000GB

Class A Operations: 1000000 per Month



Class B Operations: 1000000 per Month

Hosted on Regional Storage Class: us-central1

\$25.40 per month

\$304.80 per year

- b. List the software you are recommending for this project, include programming costs if applicable.

No.	Software	Use	Cost
1	Python 	Used to code and execute the ETL Process	Free
2	Tableau Desktop 	Used to connect to data warehouse on Google Cloud for Visual Analysis and hosting of Dashboard, and users access.	Cost: Cost per user type: Tableau Creator - \$70 per year (Site Administrator, Developer) Tableau Explorer - \$42 per year (Analyst) Tableau Viewer - \$15 per year (Field User)

Web Browsers supported by Tableau

Chrome on Windows, Mac, and Android 4.4 or later

Microsoft Edge & Internet Explorer 11 (in compatibility mode) on Windows

Mozilla Firefox & Firefox ESR on Windows and Mac

Apple Safari on Mac and iOS 8.x or later

4. Testing/Training

- a. **Explain your testing plan, how long will it take, how many days before we go live should all testing and training be done, who will do the scenarios:** In order to stay within budget, Make It Better (MIB) consulting has approved to swap out resources during the testing phase in order to stay within budget. During the testing phase, one data analyst will have their hours reduced to half (developer will be re-assigned with other work through MIB internally during those hours) so that an MIB QA Test Engineer can be used as a substitute to push through the testing phase. The intent here is to keep the project within budget for resources.
- b. The testing phase will be split into two sections. First with a QA automation engineer from Make It Better Consulting (only contracted during the testing

phase) and then User Acceptance Testing by administrators with Vision Zero (to be selected by client).

- i. QA Testing Phase - The QA test plan will be written in advance of testing and will include what features should be tested, which features should not be tested, the approach to testing, item pass/fail criteria, the test deliverables, test environments, and schedule. All test cases will be written out first and then, when the system is ready, manually tested. The manual testing phase will begin with the QA reviewing the requirements and writing up all potential test cases around the system. Once the system is ready, manual testing can begin, and all failed test cases would be sent back to the developer.
- ii. UAT Testing Phase - A guideline of test cases will also be written out by QA for UAT. These test cases are mandatory test cases for users to follow to ensure that all major aspects of the system are verified by the user. The user will then test in an exploratory manner in order to add additional confidence to the system. If any features were implemented incorrectly or if any bugs are found, the developers will be notified and corrections will be made. Retests by both QA and UAT will follow.
- iii. The testing phase is scheduled to begin 1 month before launch. This would be optimal, since according to our schedule, all major features and integration should be completed and working by this point. 2 weeks will be carved out for testing, with 1 week assigned for QA testing, and 1 week for UAT. However, the ratio of time for QA/UAT can be adjusted as necessary.

c. How will the results support the reliability criteria?

- i. Testing will help to support confidence in the system. Minimum requirements for a production build are defined as “no errors or critical/major bugs”. Minor/Trivial issues will be discussed with stakeholders first, and then prioritized for highest importance. Once a stable build is created with only minor/trivial bugs remaining, developers will begin working on the highest priority issues first. All remaining minor issues can be pushed off till after launch if time constraints exist.
- ii. Testing will also help to identify the level of quality. Quality being defined as the degree to which our product has fulfilled the clients requirements. Through user acceptance testing and sign off, we can confirm that there are no surprises and that the user requirements have met expectations.

d. Explain your training plan, who will do it, for how long

- i. Business Analysts will provide product training, which will be split up in 3 parts. Training will be scheduled 1 week before the product hand off. This will ensure that the software build in use for training is production ready.
 1. Documentation - Throughout every phase of development, business analysts have worked on an evolving product manual. This document will be finalized once a production copy of the

system is ready and will be made available to the client to use for reference.

2. Face to Face Training Sessions - Working with the client, a full day training session will be scheduled within the last week before product launch. Documentation, both hardcopy and digital, will be provided to all attendees.
3. The Live training session will also be recorded and provided in webcast format to be referenced by client if necessary.
4. Additional free on call support will also be packaged as part of product support for up to 6 months after product launch. This can be used for support/training. Further support will have to be renegotiated at the time of support expiration.

5. Support Plan

The purpose of the support plan document is to help developers and maintainers understand the system requirements, tools, and resources needed for continuous maintenance and development of the system after the transition period.

The O3SL team will not be providing support in any form after the end of December 10th 2018. The clients will need to acquire both Google Cloud Platform and Tableau online service licenses to host the system's data and dashboards. They also will need a database engineer and analyst(s) to maintain, develop, and continue to modify the system and dashboards. For continuous system development and maintenance client can hire developers/analysts who has experience in the following areas:

- Python (Pandas, Numpy)
- Google Cloud Platform
- Tableau Online Development

Information listed below is necessary regarding the system development environment:

- Python <https://www.python.org/>
- Pandas <https://pandas.pydata.org/>
- Numpy <http://www.numpy.org/>
- Google Cloud <https://cloud.google.com>
- Tableau Online <https://www.tableau.com/products/cloud-bi>

The team will be responsible for the system deployment at the server provided by the client (We strongly recommend Google Cloud Platform). After the initial deployment, the team will not be responsible for any new releases of the system. However, the clients may look for developers or maintainers to support and continue modify the system.

For each release we received feedback from our clients to adjust and modify the system accordingly. In every next iteration, we fixed the defects pointed out and agreed from previous

releases and also made the necessary changes requested and agreed by our clients and the development team. As for the final release, the team will deploy the accepted version of the system to the server provided by the client.

The system requires a web-hosting service to host the application. The system relies heavily on using open source libraries and web services; therefore, future API changes or upgrades of these services might have impacts on the system.

There is no specific hardware or Special Facilities needed for maintaining the system except a computer/server with a modern OS supporting python development requirements.

Table 1: Python

Software Requirement:	Python 3.7.1
Rationale:	Used to code and execute the ETL Process
User/Operator Manual:	https://docs.python.org/3/
Availability Information:	Well supported, Open source
Cost:	Free

Table 2: Google Cloud Platform

Software Requirement:	Google Cloud Platform
Rationale:	Used to host the data in warehouse after the ETL process
User/Operator Manual:	https://cloud.google.com/docs/
Availability Information:	Infrastructure as a Service, Platform as a Service, Serverless Platform
Cost:	Storage Size: 1000GB Data Retrieval Size: 1000GB Class A Operations: 1000000 per Month Class B Operations: 1000000 per Month Hosted on Regional Storage Class: us-central1 \$25.40 per month \$304.80 per year

Table 3: Tableau Online Development

Software Requirement:	Tableau Desktop 2018.3
Rationale:	Used to connect to data warehouse on Google Cloud for Visual Analysis and hosting of Dashboard, and users access.
User/Operator Manual:	https://www.tableau.com/learn
Availability Information:	Analytics Platform as a Service
Cost:	Cost per user type: Tableau Creator - \$70 per year (Site Administrator, Developer)

	Tableau Explorer - \$42 per year (Analyst) Tableau Viewer - \$15 per year (Field User)
--	---

*All Costs are indefinite until the Platform is terminated or migrated to a different infrastructure

Cost formula (per year):

\$304.80 + \$70(# of Tableau Creators) + \$42(# of Tableau Explorers) + \$15(# of Tableau Viewers)

- **Tableau Creators** are the Developers that will maintain the Tableau Environment
- **Tableau Explorers** are the Analysts that will create and modify the Dashboard(s)
- **Tableau Viewers** are the field users that will physically move the cameras

Recommended Scenario:

\$304.80 + \$70(1) + \$42(5) + \$15(20) = \$884.80 per year

6. Implementation Plan

- a. Describe your implementation plan, which method will you use and why?
 - i. Since there is no current system being replaced, Go-Live will occur as soon as testing and training is complete. We will not have to cut over from a previous system, run our system in parallel with an older one, or worry about adjustment. The O3SL dashboard is a new and unprecedented addition to the tools the NYPD has at its disposal. Therefore, go-live will provide quick and immediate value without interrupting previous processes. They can gradually start using it without risk of lost productivity. December 10th is the Production go-live date.
 - ii. We will have a 1 month run-up window that includes QA testing, UAT testing and on-site training. After this, the production suite will be enabled and handed over to the NYPD as fully functional. User accounts will be created and used as required and by Vision Zero team. No further support is provided at the included price after go-live.

7. Variance from Original Specifications

- a. Deviations from the original design
 - i. The original discussions called for fixed camera locations. After the Project Charter but before the Design Phase, based on feedback, it was decided that we would account for movable cameras. This increased our scope as the tool would not be used just once, but can be reused multiple times. When the number of violators caught by a specific camera decrease past a certain threshold, the camera can be reassigned to the next dangerous location. This increased the scope of our project, but it did not increase costs since it was done so early in the project. The analysts were able to accomplish the additional scope within the same development time originally estimated. It also saves the NYPD camera

costs since they will be reusing the same cameras.

8. Proposed Modifications

- a. Updating the Critical Score Formula: The recommendation system used by O3SL was created with a weight formula using accidents and fatalities. The weight for each was provided to O3SL developers by the NYPD during data-gathering. A possible enhancement would be to add in the factor of time to the weight. For accidents that occurred more recently, the weight should be increased. If multiple accidents happen in a certain area, the weight for that accident should shift appropriately. By taking time into account, cameras can be added to areas where there may be an uptick in recent accidents.
- b. Add Automatic Zoom: Currently, users have to manually zoom into a location on the dashboard map. By adding an automatic zoom feature, when users click on a map point, the map would zoom into a comfortable street level view.
- c. Add Google Street View: By adding google street view to the maps, users would be able to click on a location, and explore the actual street area. Street view would also allow to explore adjacent areas. This level of detail may help add additional insight even before visiting the potential location for a mobile camera assignment.

9. Decision Log

ID	Decision	Description of decision and impact	Source	Owner	Priority	Date Added	Status	Decision Date	Record of decision made
Unique identifier	Brief 3-5 word description of decision needed	Description of the decision and the impact of making or not making the decision	Meeting/conversation/person from which this need for a decision came	Person who is responsible for ensuring this decision is made	Relative priority: 1-High, 2-Medium, 3-Low	Date this decision was added to this list	Status of item: 1-Not started, 2-In progress, 3-Complete	Date this item was complete.	Notes or results on the completion of this item, such as final decision document
4A	Decide Document and Presentation Responsibilities	Distribute report and presentation duties. This was needed in order to proceed with the project.	Group Meeting 1	Team	Medium	3-Dec	Complete	3-Dec	Implementation Phase Meeting Log
4B	Decide Presentation Flow	Discuss Length of time, what will be presented, how the demo will look, which features will be shown.	Group Meeting 1	Team	Low	3-Dec	Complete	3-Dec	Implementation Phase Presentation slides
4C	How to resolve account expiry	We ran into an issue where the Tableau account that was prepared for the Phase IV demonstration was set to expire on the day of the presentation. We had to decide either to 1. pay for Tableau or 2. move the entire project to a different account to take advantage of another free trial. We decided to go with option 2.	Group Meeting 2	John	High	6-Dec	Complete	6-Dec	New working Tableau Account

10.Appraisal Sheet

Group Meeting Log Sheet 1

Date of Meeting: 12/03/2018

Time of Meeting: 7:00PM

Group: Make It Better

Recorder: Sage Ashique

Attending:

- Sage Ashique
- John Ortiz
- Eren Yuceer
- Suhaib Alsahybi
- Paridhi Sharma

Absent: - None Excused?

Meeting Minutes

- Divide Work for Phase 4 Paper
- Discuss Final Updates for O3SL Product
- Divide Work for Phase 4 Presentation
- Test Final version of Dashboard individually
- Discuss Lessons Learned

Tasks Assigned	Team Member	Delivery Date
Support Plan / Final Updates to Dashboard Code	John Ortiz	12/05/2018
Executive Summary / Testing/Training / Proposed Modifications	Sage Ashique	12/05/2018
Project Summary	Eren Yuceer	12/05/2018
Hardware/Software	Paridhi Sharma	12/05/2018
Implementation Plan, Variance Section	Suhaib Alsahybi	12/05/2018

Performance Appraisal & Sign-off

Meeting Ending Time: 9:30 PM

Team member name	Signature	Contribution
John Ortiz		20%
Sage Ashique		20%
Eren Yuceer		20%
Paridhi Sharma		20%
Suhaib Alsahybi		20%

Group Meeting Log Sheet 2

Date of Meeting: 12/06/2018

Time of Meeting: 7:00PM

Group: Make It Better

Recorder: Sage Ashique

Attending:

- Sage Ashique
- Eren Yuceer
- Paridhi Sharma
- John Ortiz
- Suhaib Alsahybi

Absent: - None Excused?

Meeting Minutes

- Finalize Phase 4 Paper
- Work on live presentation
- Finalize Powerpoint presentation
- Rehearse live presentation
- Agree on Presentation flow
- Discuss expiring license issue

Tasks Assigned	Team Member	Delivery Date
Prepare O3SL Dashboard	John Ortiz	12/08/2018
Review Entire Paper	Sage Ashique	12/08/2018
Add details from previous Phases into final paper	Eren Yuceer	12/08/2018
Finalize Presentation Slides	Paridhi Sharma	12/08/2018
Prepare Presentation Flow	Suhaib Alsahybi	12/08/2018

Performance Appraisal & Sign-off

Meeting Ending Time: 9:00 PM

Team member name	Signature	Contribution
John Ortiz	<hr/>	20%
Sage Ashique	<hr/>	20%
Eren Yuceer	<hr/>	20%
Paridhi Sharma	<hr/>	20%
Suhaib Alsahybi	<hr/>	20%