Analysis of Red Pill Analytics Data

Using Tableau

Business Intelligence and Analytics Group Project Summary Report

Group Members

- Group #14
- Shrikant Tambe
- Sagar Vasekar
- Shekhar Bhosale

Red Pill Analytics using Tableau

David Eccles School of Business

Table of Contents

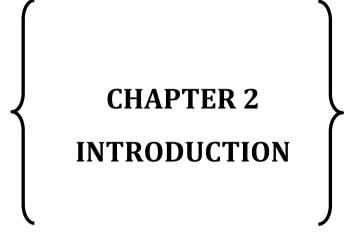# CHAPTER 1
EXECUTIVE SUMMARY

1. Executive Summary

Modern IT world is taking shift fast as blink of eyes! One of the evolving factor is project management and planning. Project management is a growing field used increasingly by businesses of all sizes. As entrepreneurs and company executives deal with the daily responsibilities of managing an organization, it is important to use dedicated project managers to oversee projects from conception to completion.

Agile methodologies proved itself as one of the effective way of project management. Agile management, or agile process management, or simply agile refers to an iterative, incremental method of managing the design and build activities of engineering, information technology and other business areas that aim to provide new product or service development in a highly flexible and interactive manner; an example is its application in Scrum, an original form of agile software development. The Agile Manifesto is centered on four values: communication with parties is more important than standard procedures and tools, focus on delivering a working application and less focus on providing thorough documentation, collaborate more with clients, and last be open to changes instead of freezing the scope of the work. However, many organizations choose to use tool to keep track of this whole project management activities. JIRA is an issue management platform that allows teams to easily manage their issues throughout their entire lifecycle. It is highly customizable and can be tailored to fit any workflow you need. JIRA is a software tool used to manage development projects that utilize the Agile methodology. It also allows to extract data from backend using its APIs.

With an advancement of Business Intelligence and analytics, we can fetch some trends and metrics to understand how project management going so far. Few of the use cases would include, to check productivity of newly hired resources, to see if project is on track in terms of commitment issues, to check closer rate of issues, to check sprint burndown charts, to check doneness of user stories, etc. These analytics could help management take future actions depending on current state or historical trends.

Our analytics for Red Pills is aimed to find certain questions stakeholders had pertaining to usage of JIRA and issues. With the help of data analysis, we aimed three main target areas – Timesheets, Issues and project analytics. Whole idea is clean and analyze data to fetch important metrics that would answer questions related to timesheet details, issues trends, sprint workload and key project metrics. With that, we started with cleaning data first and arranging data in logical orders. With collaboration of Red Pills point of contact, we worked on three

stories that dictates few observations. Suggested changes were implemented in charts after talking with company official as well. At the end of this project, we could come up with few suggestions and recommendations for Red Pills Analytics project management team. Please read through this report for details. We strongly believe that that would prove answers to questions Red Pills was seeking.

CHAPTER 2

INTRODUCTION

2. Introduction:

2.1 Red Pill Analytics

Our project was associated with a company called Red Pill Analytics. They basically wanted to do Project Management analysis where they could see which users are having more workload, who is not so active within projects and which projects are taking more time etc.

Red Pill Analytics specializes in Oracle Business Intelligence, Oracle Data Integration, Oracle Analytics, Big Data, Capacity Analytics, Agile Analytics, OBIEE 12c, Training and Checkmate.

2.2 Group Details

We had worked in a group of 3 on this project. Our details are as below:

Shrikant Tambe (U1061256)

Sagar Vasekar (U1061162)

Shekhar Bhosale (U1061255)

2.3 Project Description

a. Our goal was to help Red Pill Analytics to fetch some important metrics and trends from project management tools 'JIRA'. This information could prove useful in their near future decision making.

b. Red pills have engagements which run in their own JIRA environment as well as several running in the client's environment! It is also very probable they will eventually have clients that leverage tools such as Microsoft Project, Asana, Trello, VersionOne, Mavenlink, etc. As a result, Red Pill was seeking an Analytics and Reporting solution that will be business centric and tool neutral: They were looking to be able to integrate different data sets across different engagements and different clients and analyse and track the business metrics that can be important to them, regardless of the tool the data came from.

c. Red Pill Analytics runs most of its engagements using the Scrum Methodology of Project Management. They have a custom adaptation of this methodology deemed as methodology "Capacity Analytics"! Under Capacity Analytics, clients purchase the capacity of a full-time employee they wish to engage with. Capacity Analytics permits
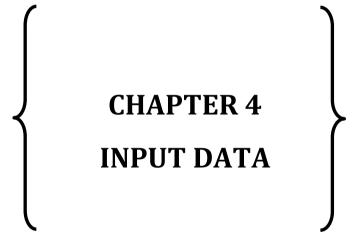
Red Pill's client to purchase 50% of a FTE's Capacity, 100%, 1.5 FTE Capacity, etc. Under Capacity Analytics, work is estimated by complexity (Story Points) vs. number of hours. In scoping by complexity vs. number of hours, they are adopting "It is better to be vaguely right than exactly wrong." Methodology! Apparently, not all clients have adopted this methodology, however, so some of their engagements utilized time tracking/timesheets and time estimates. However, Red Pills Analytics also offer definitive, pre-scoped consulting services. As a result, they are seeking an Analytics and Reporting solution that can translate traditional work and estimates into less traditional approaches, estimates and measurements. This translation would happen on the technical Red Pill side and be delivered already completed to the Analytics and Reporting environment for consumption. Red Pill Analytics is seeking a reporting and analytics solution that will provide insight to how engagements are running like "How much time/complexity are they taking vs. what was estimated?", "What are workloads/completed work by employee?" Etc.

CHAPTER 3

PROBLEM STATEMENT

3. Problem Statement

Red Pill Analytics wanted to do Project Management analytics. Following are some of the use cases or question they were more concerned with:

a. How many issues are there, on average? By Sprint/Engagement/Client/Assignee/Status/ Complexity (Story Points)/Assignee by Status/Assignee by Complexity/Assignee by Status and Complexity?

b. What is the outstanding (in-progress, not yet started but in open Sprint) workload (complexity and/or work time estimate) per assignee?

c. What is the workload (completed, in progress, and not yet started) breakdown (bucketed by assignee) per Sprint? Workload bucket size is based on complexity (story points) and/or work time estimate.

d. Get Engagement allocation (By Client/avg. duration/Project Management Env. and Tool),

e. Traditional timesheet reporting (By Engagement/Client/Worker/Worker by Engagement per Client, Average/Median/Mode), Issue Open Time (Closed - Started) vs. Complexity

f. Where does is appear, issues are/are not closing timely?

g. How many issues are being updated/kept up to date as they are being worked vs. how many are being "bulk entered/updated"?

h. How many Sprints are being planned/put into the tool after Sprint has already "started"? (issue entered dates/time > Sprint start date/time)

i. What is the average activity by issue/person? (Comments, Worklogs, etc.)

j. What are the issue with more number of Comments?

k. Are some people entering comments more than others?

CHAPTER 4
INPUT DATA

4. Input Data

a. Data given to us by Red Pill was spread across 6 worksheets. Each sheet comprised of different details.

b. **Sheet 1(Projects):**

    i. 1st sheet comprised of Project details. It tells about the projects undertaken by Red Pill.

    ii. Table structure for this sheet was as below:

Project ID	Project Key	Project Name	Project Category	Project Category Description

    iii. Each of these columns gave following information:

        1. Project Id: Unique identifier for that project.

        2. Project Key: Short Key used to refer that project.

        3. Project Name: Name of the project.

        4. Project Category: Category to which that particular project belongs.

        5. Project Category Information: Information about that category of that project.

c. **Sheet 2(People):**

    i. 2nd sheet gave us information about the employees of Red Pill Analytics.

    ii. Table structure for this sheet was as below:

User Key	User Name	User Email	User Status

    iii. Each of these columns gave following information:

        1. User Key: Short key used to refer that user.

        2. User Name: Name of that user.

        3. User Email: Email id of the user.

        4. User Status: Active/Inactive status of that user.

d. Sheet 3(Issues):

   i. 3rd sheet was used within Red Pill Analytics to keep log of issues occurred across different project.

   ii. This table contained following columns:

   1. *Issues ID:* It stored the unique id of each issue logged.

   2. *Project Key:* It tells us about the project in which issues has occurred.

   3. *Issue Key:* Unique identifier for the issue within that particular project.

   4. *Issue Summary:* Summary of issue.

   5. *Issue Description:* Detailed description of issue.

   6. *Parent Issue ID:* If it's a child issue which is dependent on any other parent issue then Parent Issue ID tell about that parent issue.

   7. *Issue Type:* Tells us about the type of that particular issue.

   8. *Issue Status:* What's the current status of issue?

   9. *Issue Assignee (User ID):* User to whom this issue is assigned.

   10. *Issue Reporter:* User who has reported this issue.

   11. *Issue Creator:* User who has created the issue.

   12. *Issue Created:* Data when issue was created.

   13. *Issue Last Updated:* Date and Time when the issue was last updated.

   14. *Issue Last Viewed:* Date and Time when the issue was last viewed.

   15. *Due Date:* Deadline for the issue resolution.

   16. *Sprint:* Sprint in which issue has occurred.

   17. *Story Points:* Points for the story.

e. Sheet 4(Work logs):

    i. This sheet gave information about the work logs created in Red Pill Analytics corresponding to issues mentioned in issues sheet.

    ii. Table structure for this sheet was as below:

Issue ID	Work Log Number	Work Log Entry Comment	Work Log Entry Date/Time	Work Log Entry Worker	Work Log Entry Duration (seconds)

    iii. Each of these columns gave following information:

        1. Issue ID: Issue ID for that issues taken from issues sheet.

        2. Work Log Number: Unique identifier to identify each work log.

        3. Work Log Entry Comment: Comment made on the work log.

        4. Work Log Entry Date/Time: Date and Time when the work log entry was made.

        5. Work Log Entry Worker: Person who is going to work on the work log.

        6. Work Log Entry Duration(seconds): Duration of particular work log in seconds.

f. Sheet 5(Attachments):

  i. This sheet gives information about the attachment made on comments and work logs.

  ii. Attachments can any different type of document which helps in resolving the issue.

  iii. Table structure for this sheet was as below:

Issue ID	Attachment Number	Attachment Date/Time	Attacher	Attachment Name/File	Attachment Link

  iv. Each of these columns gave following information:

  1. Issue ID: Unique issue id taken from issues worksheet.

  2. Attachment Number: Unique identifier for that attachment.

  3. Attachment Date/Time: Date and Time when Attachment was made.

  4. Attacher: Person which made that attachment.

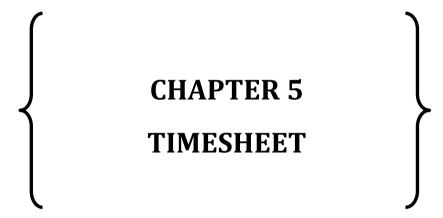 *5. Attachment Name/File:* File Name.

6. *Attachment Link:* Link for that attachment.

g. Sheet 6(Comments):

i. This sheet tells about the comments made on any particular issue.

ii. Table structure for this sheet looks like below:

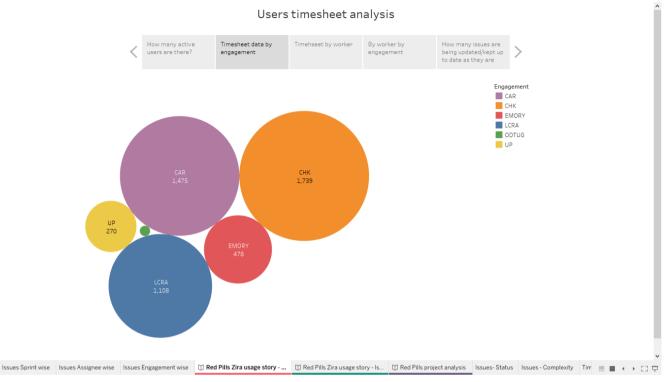| Issue ID	Comment Number	Comment Date/Time	Commenter	Comment Text

iii. Each of these columns gives following information:

1. *Issue ID:* Unique issue id taken from issues worksheet.

2. *Comment Number:* Unique identifier for that comment.

3. *Comment Date/Time:* Date and Time when comment was made.

4. *Commenter:* Person who made that comment.

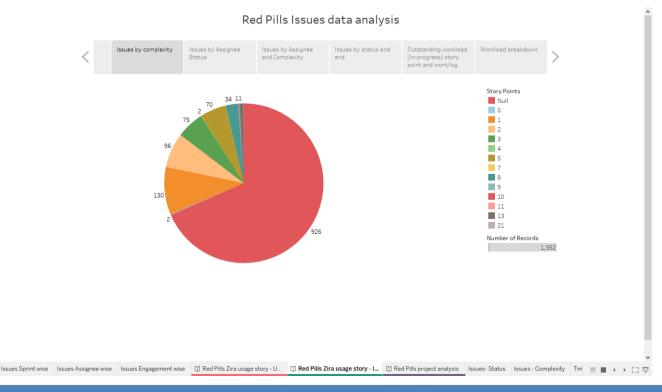5. *Comment Text:* Text of the users' comment.

CHAPTER 5
TIMESHEET

5. Story – Timesheet

In order to better organize the different visualizations and data we have divided all visualization into 3 stories as below:
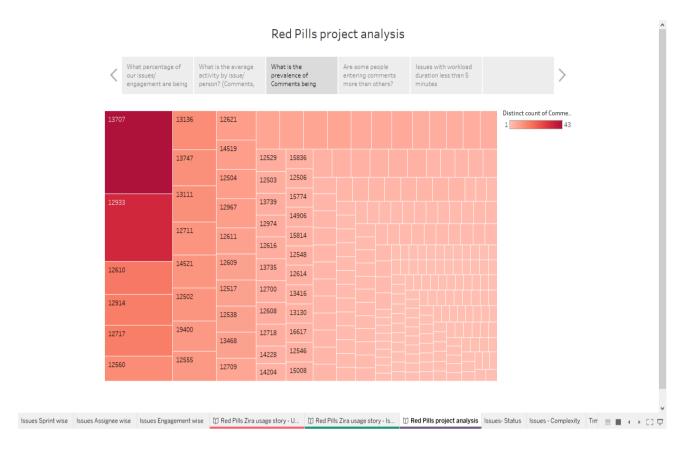
- Timesheet – This story basically deals with user related data.

- Issues – This story contains visualization related to issues raised across different projects.

✛ Project Analysis – This story comprises of visualization based on analysis of project related data of Red Pill Analytics.

Timesheet story consists of 5 different use cases as below:

✛ How users can be classified into? What are the number of internal and external users per classification?

✛ How many hours of work is done per Engagement?

✛ How many hours of work is done by each Worker?

✛ How many hours of work is done per Engagement per Worker?

✛ How many issues are being updated/kept up to date as they are being worked vs. how many are being "bulk entered/updated"?
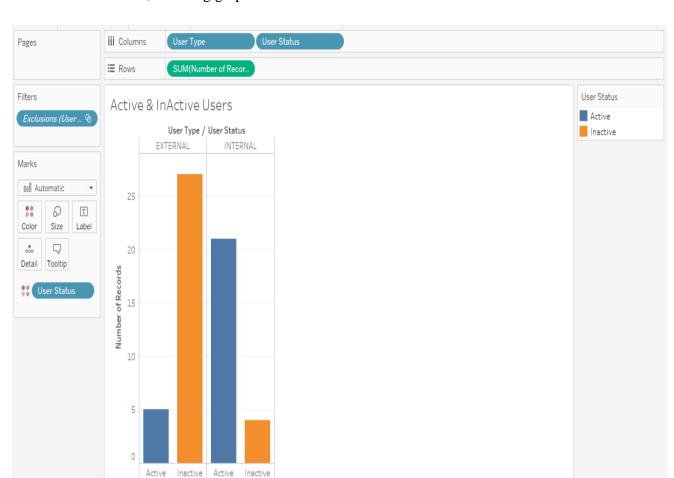
This story focuses on answering the questions related to timesheet hours spent. We extracted data from Work logs and Issues dataset to build the useful and informative results. Let's go through each use case one by one.

5.1 How users can be classified into? What are the number of internal and external users per classification?

Major challenge in building this analysis was identifying User attribute based on which users can be classified. For that, we considered Email_ID field of user. Users with Email_ID of "redpillanalytics.com" are considered as "Internal Users". Users with all other Email_IDs are considered as "External Users". We excluded users with Email_ID as "atlassian.com" from this classification.

To answer this use case, we have prepared a Vertical Bar chart. Since Vertical Bar charts are always easier to compare progressive data values for different attributes and provides generalized view. So, it easily grabs attention of viewer about how many Active/Inactive users are there in each classification of users i.e. External or Internal.

In column shelf, we took User Type and User Status and on row shelf we took SUM of number of records. For color, we have used User Status and in filter we have added an exclusion to exclude users with email id "atlassian.com"

Therefore, resulting graph looked like below:

In this way, this chart answered the use case of classifying users according to the internal and external type and also, categorizing them into active and inactive type.

5.2 How many hours of work is done per Engagement?

While analysing Timesheet related data, first question that may come across is "How many work hours are getting spent on Engagement?" To answer this use case, we have prepared a Packed Bubble chart. Since Packed Bubble charts represents data in such a way that entities with higher data value are displayed with big bubbles. So, it easily grabs attention of viewer about which Engagement has highest number of work hours spent and which has lowest work hours spent. Therefore, we had used packed bubble chart for this question.

We used data about Work Logs of workers and data about Issues. Work Log Entry Duration in Work Logs dataset, represents duration taken on a given issue by a particular

worker in seconds. Data in seconds was complicated to understand as it became huge when summed up across for entire Engagement. So, we created calculated field named "Work Log Entry Duration(In Hours)". We converted this field to measure.
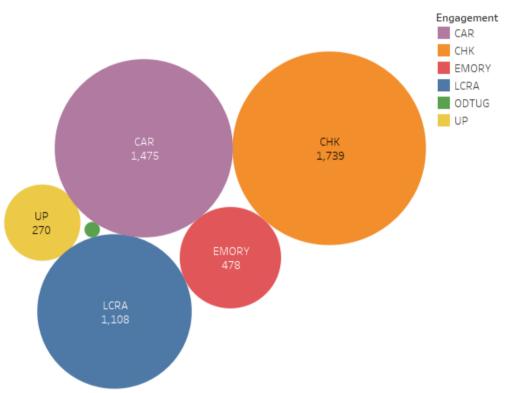
Work Log Entry Duration(In Hours)=[Work Log Entry Duration(In Seconds)]/3600

For Color, we took "Engagement".

For Size, we took newly created measure "SUM(Work Log Entry Duration(In Hours))".

On Label, we took "SUM(Work Log Entry Duration(In Hours))" and "Engagement".

Therefore, resulting graph looked like below:

By Engagement

Engagement and sum of Work Log Entry Duration(In Hours). Color shows details about Engagement. Size shows sum of Work Log Entry Duration(In Hours). The marks are labeled by Engagement and sum of Work Log Entry Duration(In Hours).

As we can see in above graph, "CHK" engagement has maximum number of work hours utilized which are equal to 1,739 hours. Whereas "ODUTG" engagement has minimum number of work hours spent which equals 11 hours.

5.3 How many hours of work is done by each Worker?

It is one of the vital use case as it plays important role in deciding performance of individual worker. To answer this use case, we have prepared a Horizontal Bar chart. Since Horizontal Bar charts are always easier to compare data values for different attributes and provides generalized view. So, it easily grabs attention of viewer about which Worker has highest number of work hours spent and which has lowest work hours spent.

We used data about Work Logs of workers and data about Issues. Work Log Entry Duration in Work Logs dataset, represents duration taken on a given issue by a particular worker in seconds. So, we used calculated field created in earlier chart named "Work Log Entry Duration(In Hours)".

Work Log Entry Duration(In Hours)=[Work Log Entry Duration(In Seconds)]/3600

In Measure(Column), we took SUM of Work Log Entry Duration(In Hours).

In Dimension(Rows), we took "Work Log Entry Worker".

On Label, we took "SUM(Work Log Entry Duration(In Hours))".

Below is the resulting graph:

As we can see in above chart, "Stewart" is the only resource who has work hours around 1700 hours. We can conclude that Stewart has maximum work load.

5.4 How many hours of work is done per Engagement per Worker?

For a broader picture, it is important to know how work hours are distributed within given engagement across different workers. To answer this use case, we have prepared a Horizontal Bar chart. Since Horizontal Bar charts are always easier to compare data values for different attributes and provides generalized view. So, it easily grabs attention of viewer about which Worker in which Engagement has highest number of work hours spent and which has lowest work hours spent.

We used data about Work Logs of workers and data about Issues. Work Log Entry Duration in Work Logs dataset, represents duration taken on a given issue by a particular worker in seconds. So, we used calculated field created in earlier chart named "Work Log Entry Duration(In Hours)".
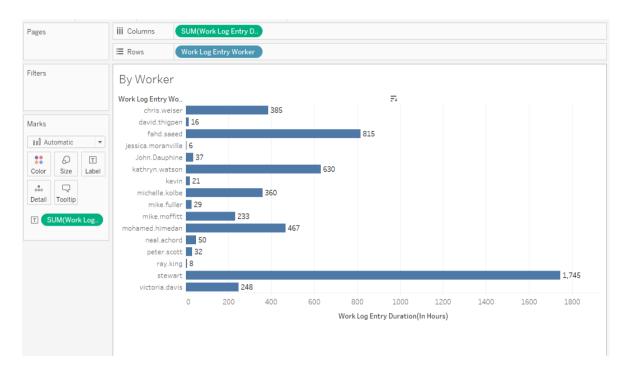
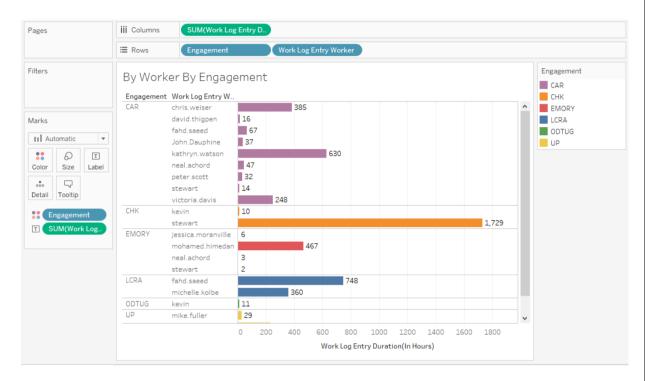Work Log Entry Duration(In Hours)=[Work Log Entry Duration(In Seconds)]/3600

In Measure(Column), we took SUM of Work Log Entry Duration(In Hours).

In Dimension(Rows), we took "Work Log Entry Worker" and "Engagement".

On Label, we took "SUM(Work Log Entry Duration(In Hours))".

On Color, we took "Engagement".

Below is the resulting graph:

As we can see in above graph, work load needs to be divided equally in Engagement "CHK" as worker: "Stewart" is over-utilized and worker: "Kevin" is under-utilized.

5.5 How many issues are being updated/kept up to date as they are being worked vs. how many are being "bulk entered/updated"?

After developing timesheet reports, one of the important use case arises that demands data for how many issues are updated right on time vs bulk update. To answer this use case, we have prepared a Vertical Bar chart. Although chart very simple, it has long calculation behind that.

We used data about Work Logs of workers and data about Issues. Work Logs table has "Work Log Entry Date\Time" field which represents time when worker worked on particular Issue. Now, in Issues dataset, there is a field named "Issue Last Updated" which represents time when Issue was updated. So, to find out Update status of any issue, we used these 2 fields and created one calculated field named "update_status".

Update_status= IF ([Issue Last Updated]<[Work Log Entry Date/Time]) THEN "Not Updated" ELSE "Updated" END
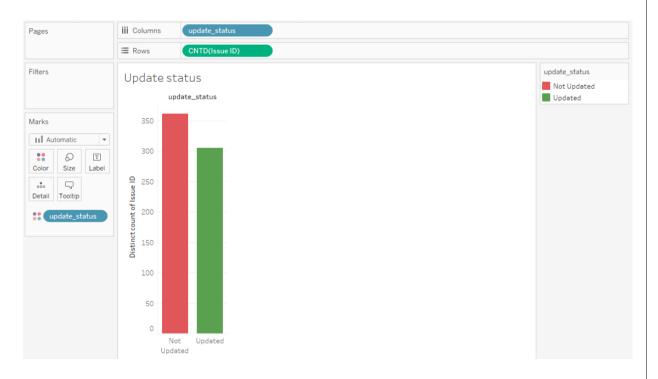
Based on this Update_status field, we figured out how many issues are updated on time and how many are bulk updated.

In Measure(Rows), we took DISTINCT COUNT of Issue ID

In Dimension(Columns), we took newly created field "update_status".

On Color, we took "update_status".

Below is the resulting graph:

As we can see in above graph, total 305 issues are Updated on time and 361 are bulk update or not updated on time.

CHAPTER 6

ISSUES

6. Story – Issues

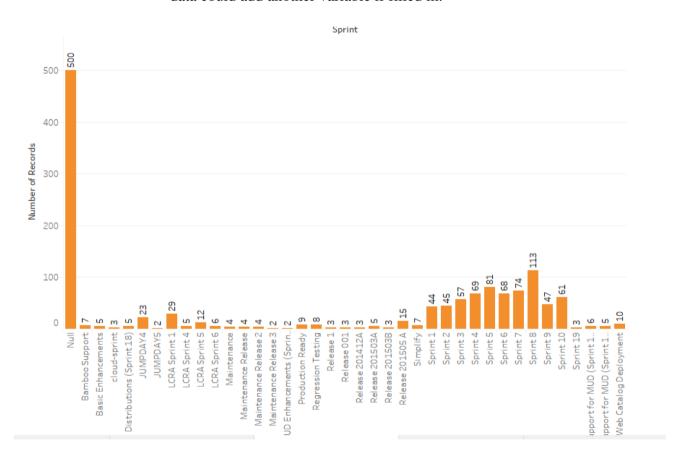**This story consists use cases related to issues as described below:**

- How many issues are there, on average? Show different charts explaining issues status with different parameters like,

    - by Sprint?

    - by Engagement?

    - by Assignee?

    - by Status?

    - by Complexity? (Story Points)

    - by Assignee by Status?

    - by Assignee by Complexity?

    - by Assignee by Status and Complexity?

- What is the outstanding (in-progress, not yet started but in open Sprint) workload (complexity and/or work time estimate) per assignee?

- What is the workload (completed, in progress, and not yet started) breakdown (bucketed by assignee) per Sprint? Workload bucket size is based on complexity (story points)

Major focus on these reports is issues and workload details. This data contains thousands of issues with history that gave us good base to start with. This huge data helps see history trends clearly. Information painted considers various factors and data associated with each issue.

6.1 How many issues are there, on average? Show different charts explaining issues status with different parameters.
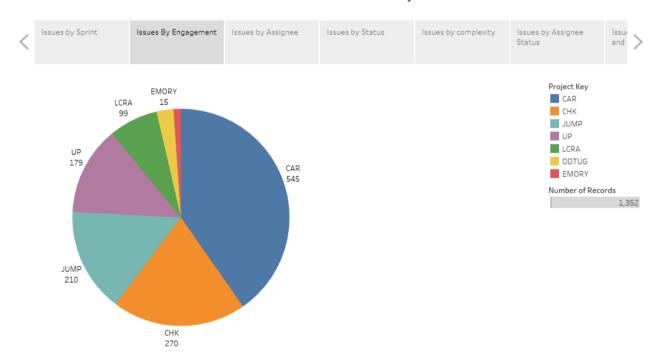
These graphs query issues table and gives us below visualization as per parameters specified.

- Issues by Sprint
  - This visual representation is speaking about issues sliced by "Sprints".
  - This graph tells us that Sprint 8 was most active in terms of issues logged followed by sprint 5
  - We have major chunk of issues (500) with no sprint mentioned. That does data could add another variable if filled in.
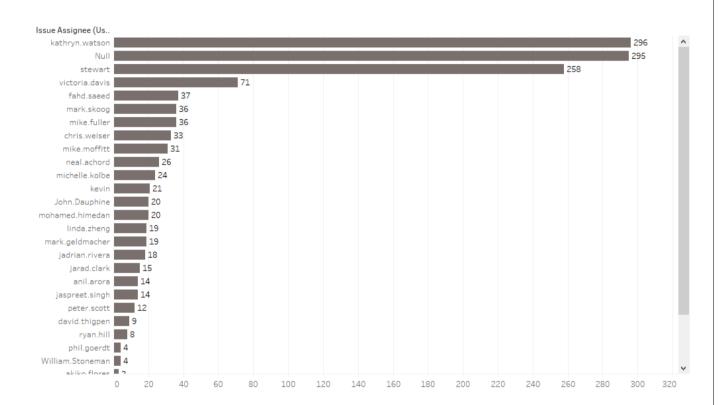
- Issues by Engagement
  - Engagements are nothing but projects.
  - Below pie chart divides issues per Projects. This was we can tell which project is more active in terms of issues.
  - This visualization shows project wise issues logged. This pie chart tells us CAR has maximum number of issues logged.
  - We have used different colours to indicate different projects issues.

Red Pills Issues data analysis
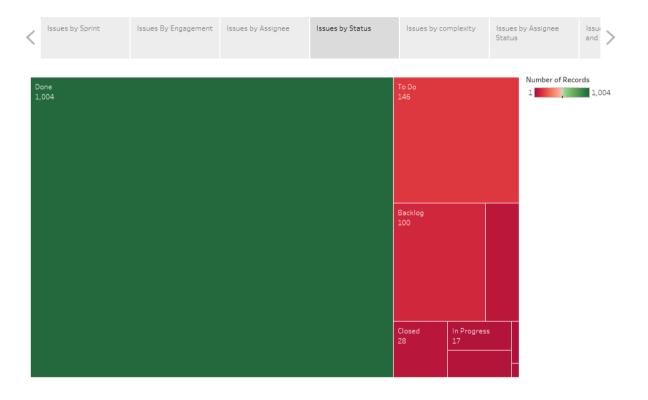
- Issues by Assignee
  - Below representation shows issues divided by "Assignee".
  - This visualization shows number of issues logged by assignee. We can see that Kathryn has more issues as compared with others.
  - We can see there are almost 300 issues with no Assignee. Red pills should fill in this data so that they do not lose track of issues because of absence of issue owner.
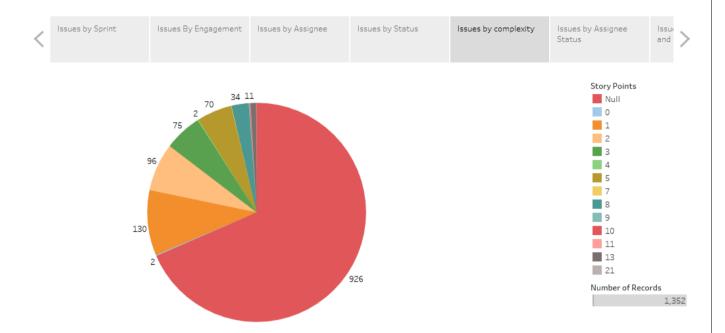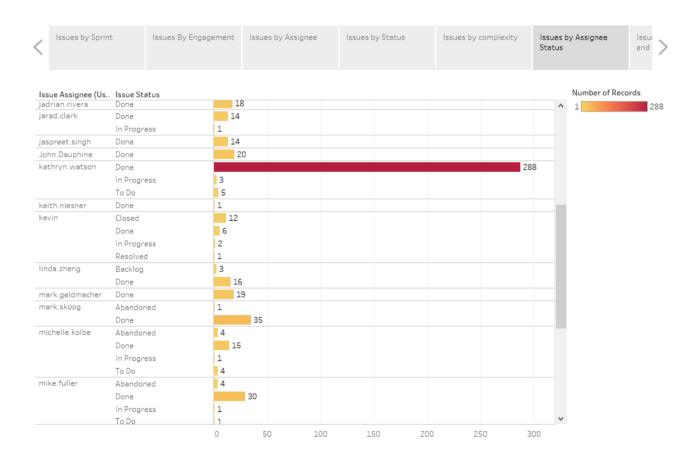
- Issues by Status
    - This visualization shows issues closure rate is pretty much high. WE have only 104 issues pending to be fixed which are targeted.

| Issues by Sprint | Issues By Engagement | Issues by Assignee | Issues by Status | Issues by complexity | Issues by Assignee Status | Issu and |
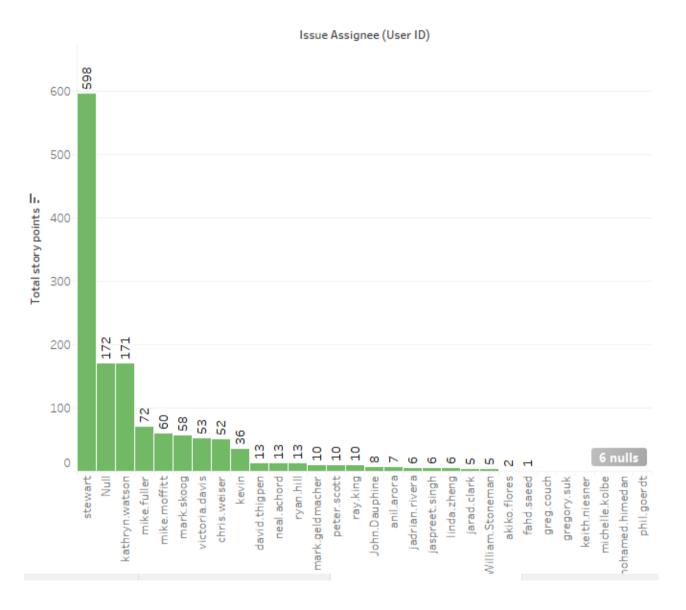
- Issues by Complexity
  - Complexity is measured in terms of story points in JIRA. Big red chunk indicated we have major issues whose story points are not specified
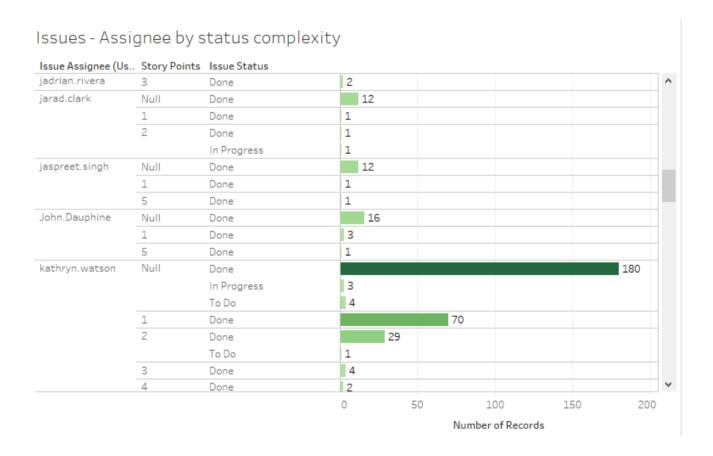
- Issues by Assignee and Status
  - Representation showed below considers more than one dimensions.
  - First we are dividing issues by assignee and second order is issue status. This way we know two level of information and it gives us more idea of current state of issues tracking.
  - This visualization further decomposes assignee bucket by issue status. We can see that Kathryn did good job in closing maximum number of issues.

- Issues by Assignee and Complexity
  - This visualization is sorted descending to indicate who did exceptional job in taking care of more complex issues.
  - Graph is fashioned in descending order to tell us where big chunk of data goes.
  - With further observation seems, Stewart handled almost 600 story points. This includes current and past assignments.

Issue Assignee (User ID)
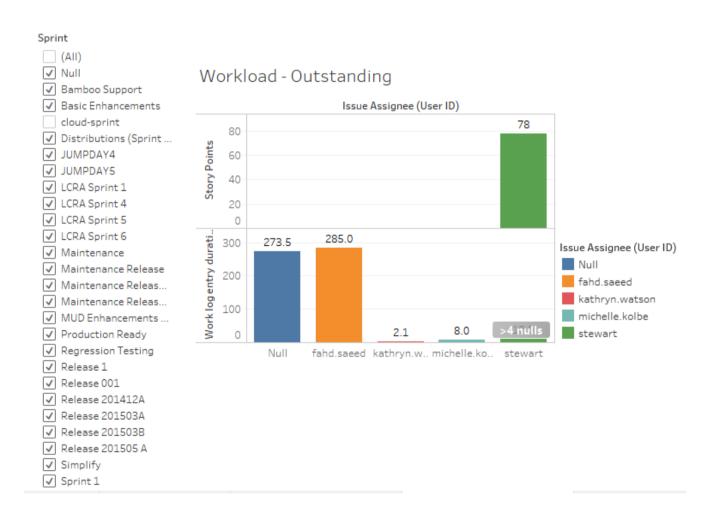
- Issues by Assignee, status and Complexity
  - To dig down little deeper, we added another dimension to assignee and status issues report: complexity. This shows three view calculation of issues.
  - With that data, we can observe that except Kathryn other users are acting on issues at average rate.
  - Most of the users have high "Done" issues. This tells, those users take more issues to close than keeping them pending. Work progress is on track or at better stage.

Issues - Assignee by status complexity

Issue Assignee (Us..	Story Points	Issue Status	Number of Records
jadrian.rivera	3	Done	2
jarad.clark	Null	Done	12
	1	Done	1
	2	Done	1
		In Progress	1
jaspreet.singh	Null	Done	12
	1	Done	1
	5	Done	1
John.Dauphine	Null	Done	16
	1	Done	3
	5	Done	1
kathryn.watson	Null	Done	180
		In Progress	3
		To Do	4
	1	Done	70
	2	Done	29
		To Do	1
	3	Done	4
	4	Done	2
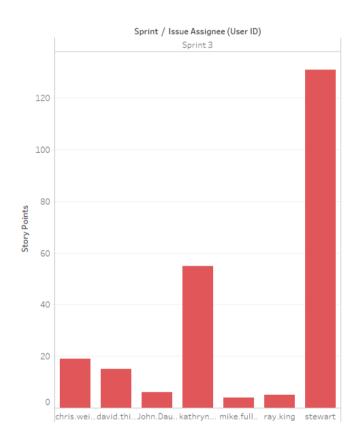
6.2 What is the outstanding (in-progress, not yet started but in open Sprint) workload (complexity and/or work time estimate) per assignee?
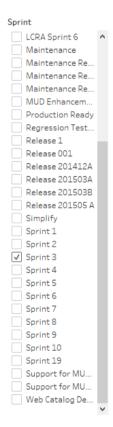
- Below graph helps answer current workload question. This bar graph is created to depicts outstanding workload of active resources. They are bifurcated by story points and work log entry duration.

- If we analyse story point wise workload, we can easily tell that there is only one user preoccupied with loads of story point tasks. We can try to even out the work and reassign some to other users so that there is not burden on one person

- If we analyse work log duration wise workload data, we can see Fahd is loaded with big chunk of work. Also, there is big block of work unassigned. We can try to assign some of unassigned work to users who do not have much on their plate.

6.3 What is the workload (completed, in progress, and not yet started) breakdown (bucketed by assignee) per Sprint? Workload bucket size is based on complexity (story points)

- Below graph is showing workload breakdown. We considered only completed, in progress and not yet started status issues to create this graph. This way we have more clear picture of what is more important.
- We have added filter for selecting desire sprint. This way we can see data about specific sprint if we want. By default, this graph will consider all sprints.
- For example, below snippet give us more details about "Sprint 3".
- With further analysis, we can mention that Stewart contributed more than others to get things done.

CHAPTER 7

PROJECT

ANALYSIS

7. Story – Project Analysis

This story consists of 5 different use cases as below:

- What issues/engagement are being actively worked/documented in the JIRA environment? (lots of comments, worklogs, etc.)

- What is the activity by issue/person? (Comments)

- What are the issue with more number of Comments?

- Are some people entering comments more than others?

- Issues with workload duration less than 5 minutes

This story focuses on answering the questions related to different projects. One of the biggest challenge for us in analysing and using the data was that, given data was not consistent across projects. We tried our best to extract the maximum actionable information from the available data. Let's go through each use case one by one.

7.1 What percent of our issues are actively worked on?

To answer this use case, we have prepared a pie chart. Since Pie charts are used to show relative proportions – or percentages – of information. In this use case, we are going to show the proportion and pie chart best suits the condition. Therefore, we had used pie chart for this question.

For this question, it was required to calculate the total number of comments and work logs for the particular Issue. So, we created one calculated field called "Comments and Worklogs" which calculated the total number of comments and worklogs for that issue. Then, created new measure to show the percentage.

On Angle, we took calculated field "Comments & Worklogs" which contained total number of comments and worklogs for that issue.

On Label, we took Issue ID.
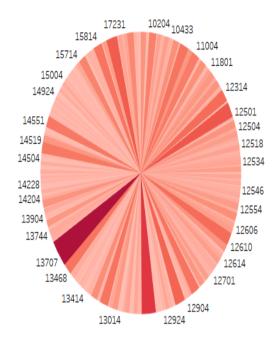
For color, we took "Comments & Worklogs".

For Details, we took newly created measure Percentage.

Therefore, resulting graph looked like below:
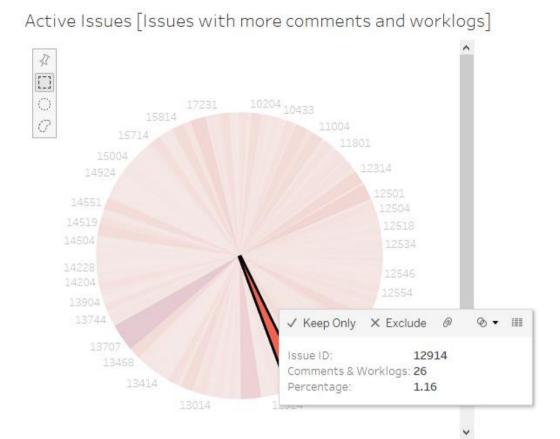
Red Pills project analysis

What percentage of our issues/ engagement are being	What is the average activity by issue/ person? (Comments,	What is the prevalence of Comments being	Are some people entering comments more than others?	Issues with workload duration less than 5 minutes

New Blank Point

Duplicate

Comments & Worklogs

2 ▭ 51

As you can see from above graph, it is clear that some issues like 13707 are very active. It is most commented issues and also there are more worklogs on this issue.

On selecting or hovering over any of the issue you can see the percentage of activeness of that issue as compared to all issue.

Active Issues [Issues with more comments and worklogs]
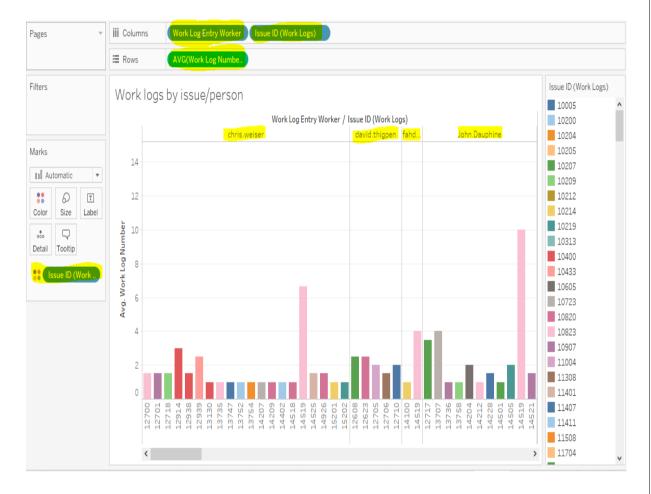
In this way, above pie chart answers the question of calculating the percentage activeness of particular issue.

7.2 What is the average activity by issue/person? (Comments, Worklogs, etc.)?

Here it was required to calculate the average activity by issue per person. Red Pill wanted to consider this use case in order to track the user participation. We took Average Work Log Number on Y-axis i.e. in Row shelf and Work Log Entry Worker, Issue Id on X-axis i.e. in Columns shelf.

For coloring we have used Issue Id. Therefore, average activity can be seen per issue/person.

Here's the final line chart showing average activity by issue/person.

7.3 What are the issue with more number of Comments?
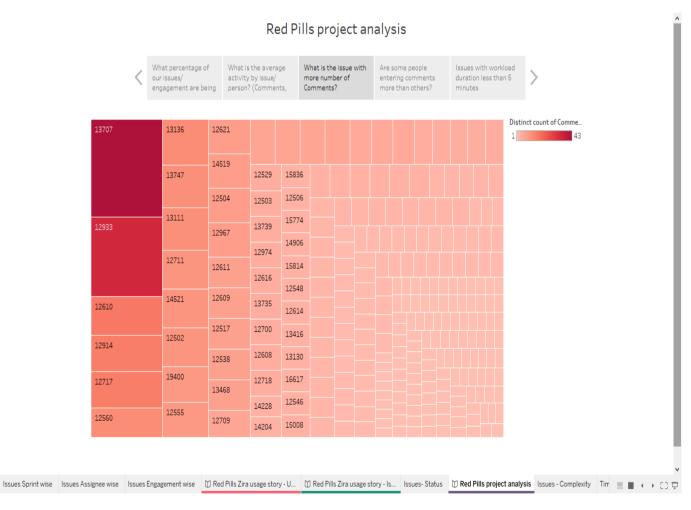
To answer this question, we have used Treemap. Treemap best suits for showing hierarchical data as a proportion of a whole. In this question, we wanted to see the issues where more comments are being entered.

Red Pill wanted to get an answer to this question, basically because they wanted to know issues with more comments. In turn, they wanted to see issues with more comments as a proportion of a whole. Therefore, we found Treemap can fit better to the situation.

For Size & Color, we took count of distinct number of comments made on issue.

For Labelling, we have used Issue Id.

Therefore, the final chart for this use case looks as below:

So, now we can clearly see, issues where the comments are more. Issue 13707 and 12933 are thus the more talked about issues.

7.4 Are some people entering comments more than others?
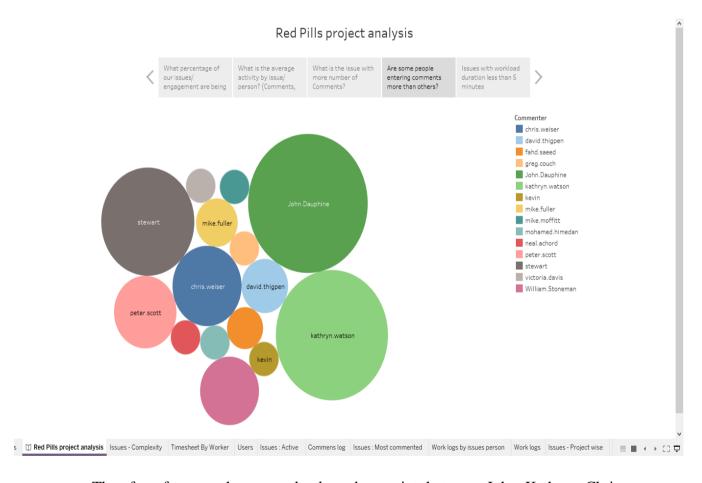
In any organization, we usually see that people who are contributing more towards helping out others are getting rewarded. For similar purpose, Red Pill wanted to know the people who have more number of comments as compared to others.

For this use case, we have Packed Bubble chart. This type of chart basically suits to the situation where we want to show the concentration of data. Here we wanted to see the users with more number of comments. Thus, this chart is the good solution for this use case.

For Label and Color, we have used Commenter.

For Size, we have used distinct number comments.

Here, the final representation answering the discussed question.

Therefore, from graph we can clearly make a point that users John, Kathryn, Chris, David, Peter and Stewart are top commenters.

7.5 Issues with workload duration less than 5 minutes
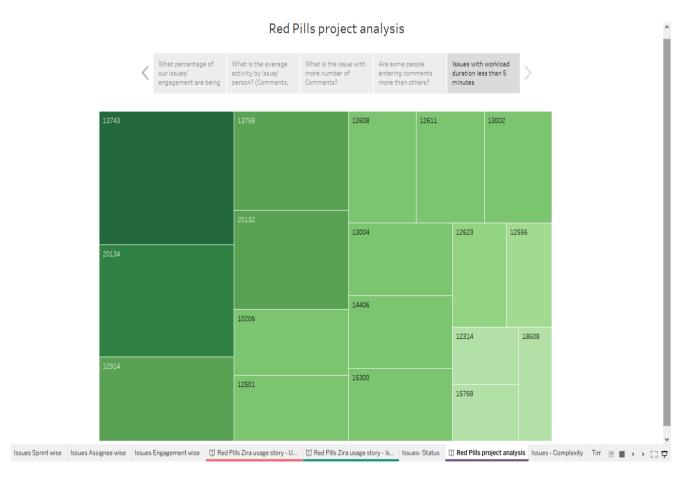
Red Pill wanted to track the issues where workload duration is less than 5 minutes. So, that they can see what's wrong with those issues and why they are having such less workload duration.

For solving this question, we have used Treemap. As discussed in few points above, Treemap can used when we want to see hierarchical data as a proportion of a whole.

Here, we wanted to see the issues with workload duration less than 5 minutes and highlight such issues. The reason why we choose Treemap here is that, we can see easily spot the issues with less duration here.

For Color and Size, we took Work log Duration where duration was less than 3600 seconds which is equals to 5 minutes for the particular issues and for label we took Issue ID.

Here's the final visualization:

Therefore, we got all our issues where duration was less than 5 minutes. Smaller the tile size lesser is the duration.

CHAPTER 8
CONCLUSION

8. Conclusion

In this way, we have prepared 3 different stories according to the type of that analysis.

Therefore, we can conclude in the same pattern based on stories.

i. Story – Timesheet

1. Maximum Time Spent on: *CAR & CHK Engagements.*

2. *"Stewart"* is the highly utilized resources among all engagements.

ii. Story – Issues

1. *"Fahd"* has maximum workload as per the work log duration & *"Stewart"* has maximum workload as per story points.

2. Sprint *"Simplify"* has maximum story points till date.

iii. Story – Project Analysis

1. Following users are more active as per the number of comments and work logs:

   Kathryn, Stewart, Chris, Peter & John

2. *18 Issues* have less than 5 minutes' duration

3. *"Chris"* has highest average number of comments/work logs per issue.

4. *Issues 13707 & 12993* are most talked about issues as per the comments made on them.

5. *Issues 13707 & 12993* are also having most number of work logs.

This analysis will definitely give some actionable information to Red Pill Analytics.

CHAPTER 9
SUGGESTIONS

9. Suggestions

a. There are only 5-6 active resources who are actively commenting on the issues. Therefore, we should encourage other team members to show active participation in issue resolution.

b. Outstanding workload indicates 2 users are loaded with high volume of work to be done and big chunk of work assigned. We could re-assign some work to those who do not have much outstanding workload.

c. We could have few more recourses on CHK project with Stewart so that Red Pills have few more people acquainted with project knowledge to reduce the absolute dependency of Stewart.

CHAPTER 10
REFERENCES

10.   REFERENCES

[1] Tableau Community - https://community.tableau.com

[2] Knowledge Base - http://kb.tableau.com/

[3] Business Intelligence Blog - http://www.tableau.com/about/blog

[4] Red Pill Analytics – http://www.redpillanalytics.com

[5] Tableau Starter Kit - http://www.tableau.com/learn/starter-kit

[6] Tableau Videos - http://www.tableau.com/learn/training

[7] IS 6481 lecture 9 - Visual Design.pdf by Michelle Kolbe

[8] Tableau Blogs

Appendix

1. Task Division

Task/Module Name	Assigned To
Story – Timesheet & all corresponding charts	Shekhar
Story – Issues & all corresponding charts	Shrikant
Story – Project Analysis & all corresponding charts	Sagar

2. Hours spent on different project tasks by each team member

Task/Module Name	Sagar [hrs.]	Shekhar [hrs.]	Shrikant[hrs]
Analysis of given data	3	3	3
Organizing Data	1	2	4
Planning and preparing the chart inputs	3	5	7
Add and maintain data connection	2	1	3
Add calculated fields, measures & dimension wherever required	1	1	1
Plotting Visualization	5	4	8
Integrating all modules	1	1	4
Creating Story	2	2	2
Preparing Presentation	1	1	1
Project Report	3	3	3

3. Use of agile project management on this project

We followed Scrum for project planning and management. We worked in two sprints. Each having three weeks of duration. Please find below details of project planning and deliveries –

Sprint   1

Duration   7-Sep to 28-Sep

Story   Red Pills analytics – Project management analysis: Requirement gathering and Data clean-up

Tasks -

Task #	Title	Story Points	Owner	Status
1	Talk to Michel about available projects	2	All	Done
2	Select project	2	All	Done
3	Follow up with company POC for project details	1	All	Done
4	Follow up with company POC for data	1	All	Done
5	Select tool	2	All	Done
8	Data clean-up: Issues table	5	Shrikant	Done
7	Data clean-up: Projects, People and comments table	5	Sagar	Done
8	Data clean-up: Worklogs and attachment table	5	Shekhar	Done

Sprint   02

Duration   29-Sep to 20-Oct

Story   Red Pills analytics – Project management analysis: Data analysis and implementation

Tasks -

Task #	Title	Story Points	Owner	Status
1	Meet with company POC (Kathryn) to understand expectations	2	All	Done
2	Plan workload and divide responsibilities	2	All	Done
3	Work on Issues story	8	Shrikant	Done
4	Work on Timesheet story	8	Shekhar	Done
5	Work on Project analysis story	8	Sagar	Done
6	Sync up with team	2	All	Done
7	Sync up with Kathryn to demo work and ask suggestions	8	All	Done
8	Implement suggestions	2	All	Done
9	Create deck and Record presentation	2	All	Done
10	Create Project report	8	All	Done

4. Feedback from your business stakeholder

Business Person Name: Kathryn Watson

As Kathryn mentioned in excel sheet, we created 10 reports for story of "Issues".

Kathryn was happy to see that our story covered all the scenarios that she wanted.

For some 2 reports, she wanted more dimensions in report. After we added those dimensions and represented data again to Kathryn. She was satisfied with the reports for entire story of "Issues" and according to her those were complete reports.

Similarly, we created 2 more stories "Timesheet" and "Project Analytics" covering all the analytical user cases provided by Kathryn.

For timesheet reports, user case #5, she suggested which columns to use to complete analysis. According to her suggestions, we made changes in Timesheet story reports and re-evaluated from her. Also, we asked for feedback for Project Analytics report. She was satisfied with reports from both the stories and asked for no corrections.

Thus, our entire analysis was approved by Kathryn and she thought it was very helpful information.

5. MoM from different project meetings with Stackholder

Meeting Date: 10th October 2016

Business Person Name: Kathryn Watson

- We discussed if Tableau is a tool to use. Kathryn is in agreement for tool

- We discussed expected outcomes. Kathryn suggested few scenarios which we could focus on

- Overall, we can slice things down project wise. Major focus would be on "CAR" project

- We shared few reports we created. Kathryn suggested few modifications to those. Those are to be implemented before we meet next time

- Kathryn suggested we could use different visualizations to showcase things instead of just bar graphs

- Plan is – Group will create reports as mentioned in information sheet and we will keep in touch with Kathryn for more suggestions.

5. MoM from different project meetings with Stackholder