Table of Contents

1.0 Summary

Sportradar is the world's leading sports technology company, at the intersection between sports, media and betting. The main focus of the Jersey color task is to add and update the jerseys for the leagues that we offer for our media and betting solutions. The idea is to ensure that the system is updated before the start date of each sport tournaments.

In this project, I overhauled our spreadsheet file and made some updates which can help my colleagues use and navigate the file more effectively. This includes a more cohesive theme, charts from pivots for quick overview of the scope, and an up-to-date data for the task.

So, the business task is, essentially, to provide my colleagues/department a more functional spreadsheet file which can be maintained and used for the succeeding years. To reach that goal, the main approach I used in this project is the same as what I've followed for my previous case study (Fitbit Data Analysis using Excel - Google Data Analytics Capstone) except for the part that this project doesn't include an analysis phase since it is not needed at all.

Following this summary each section can be viewed as one of the data analyses phases (in the same order), which commonly seen in practice. These phases are sometimes referred to as ask questions, prepare data, process/clean data, share and act.

1.1 Questions
1. How can I help optimize our workflow for the Jersey color task?
2. How can I make the spreadsheet easier to use and navigate?
3. How can I make the spreadsheet easier to maintain and update over time?

1.2 Deliverables
1. A clear summary of the business task
2. A description of all data sources used
3. Documentation of any cleaning or manipulation of data
4. Supporting visualizations

2.0 Data Preparation

In this step, we need to prepare the data for processing. We will need to take a close look at the dataset, summarize them and discover some data quality characteristics. Since we already have a dataset to work with, the first step I took is to check if we have an up-to-date data. I need to ensure that our Tour ID is the same with our system and the Start Date is correct based on our official sources. In this way, we can execute the task with ease.

2.1 Data Summary

Figure 1.0: Columns and Descriptions

Tournament	Tour ID	Priority	Level	Sport	Start Date	Countdo	Completed	Checked by	Review tim	Notes	Status
EHF Champions League	Groups	HIGH	LS1	Handball	2022-09-14	376			44818		Overdue
Chile Primera Division	67280		LCS	Soccer	2023-01-20	248			44946		Completed
Peru Primera Division	34467		LCS	Soccer	2023-01-21	247			44947		Completed
Colombia Primera A Apertura	19232		LCS	Soccer	2023-01-25	243			44951		Completed
Colombia Primera A Clasura	19236		LCS	Soccer	2023-01-26	242			44952		Completed
Paraguay Primera Division Apertura	3133		LCS	Soccer	2023-01-27	241			44953		Completed
Argentina Superliga	68		LCS	Soccer	2023-01-27	241			44953		Completed
Int. Clubs FIFA Club World Cup	3888		LCS	Soccer	2023-02-01	236			44958		Completed
Uruguay Primera Division Apertura	57776		LCS	Soccer	2023-02-04	233			44961		Overdue

Column	Description
Tournament	competition name
Tour ID	Unique ID
Priority	competitions with "High priority" should be prioritized
Level	coverage level of the competition based on offered package
Sport	Soccer, Basketball, Ice hockey, American, football, Baseball, Handball
Start Date	date when competition starts
Countdown	days until competition starts a. numbers with minus values show that the competition has not started b. numbers with plus values show that the competition has already started
Completed	date when the jersey colors for the competition were completed
Checked By	Name/Initials of the user who completed the competition
Review Time	calculated value showing how many days prior, or after, to "Start day" we completed the jersey colors
Notes	lists open topics for the competitions
Status	Current state of the competition (Overdue, Due Today, Not Started, In progress, Completed, On Hold, Cancelled)

3.0 Data Cleaning and Processing

3.1 Steps taken for data cleaning:

The following steps are done in Big Query in Microsoft Excel. Moreover, the entire project from Ask Phase to Analyze Phase was done in Microsoft Excel as I want to test my knowledge with this tool. Also, I truly believe that Microsoft Excel is sufficient for the scope of this case study.

1. Duplicate initial dataset
2. Load and Transform
3. Clean Data
   a. Duplicates
   b. Null and Blanks
   c. Invalid Values
   d. Formatting
   e. Correctness
      - Tournament
      - Tour ID
      - Level
      - Sport
      - Start Date

4.0 Visualization

We already prepared the data for visualization and processed it to make sure that it is clean. Now it is time to create a visualization to make the file easier to use and navigate. The goal of this procedure is to create 2 filters based on sport and status, 2 charts that will give an overview of the statuses (short-term and long-term overviews).

Figure 1.1: Main Header

The main header is where the magic happens. It is composed of the 4 functions mention below. Moreover, we have quick links to the Sportradar interface and manual for easy access. A color legend can also be seen on the top right for the statuses.
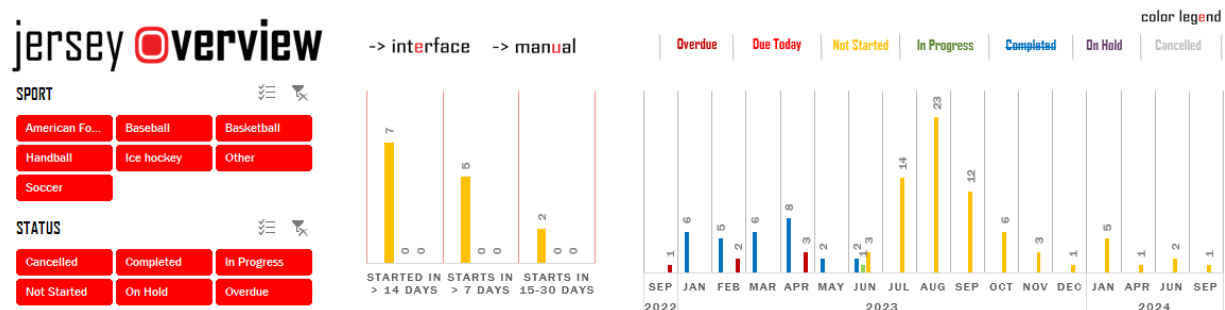
Figure 1.2: Filter by Sport

The slicer can be used to filter out rows by sport type. It can be useful when we need to narrow down our row numbers so that we can easily find specific tournaments.

Figure 1.3: Filter by Status

The slicer can be used to filter out rows by current status. It can be useful when we need to narrow down our row numbers so that we can easily find specific tournaments.

Figure 1.4: Short-term Chart

The Short-term Chart can be used to quickly identify the progress of our tournaments based on the start dates along with its status. This quick overview can be useful for knowing what and when to prioritize.

       a. Started in > 14 days (2 weeks)
       b. Starts in > 07 days (1 week)
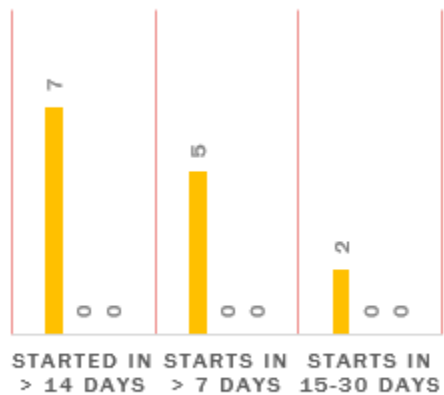   c. Starts in 15-30 days (2-4 weeks)

Figure 1.5: Short-term Pivot

The chart above is based on the Short-term pivot table below.

Days	Not started	In progress	Completed
Started in > 14 days	7	0	0
Starts in > 7 days	5	0	0
Starts in 15-30 days	2	0	0

Figure 1.6: Long-term Chart

The Long-term Chart can be used to summarize the progress of our tournaments based on the start dates along with its status. This quick overview can be useful for preparing in the future.
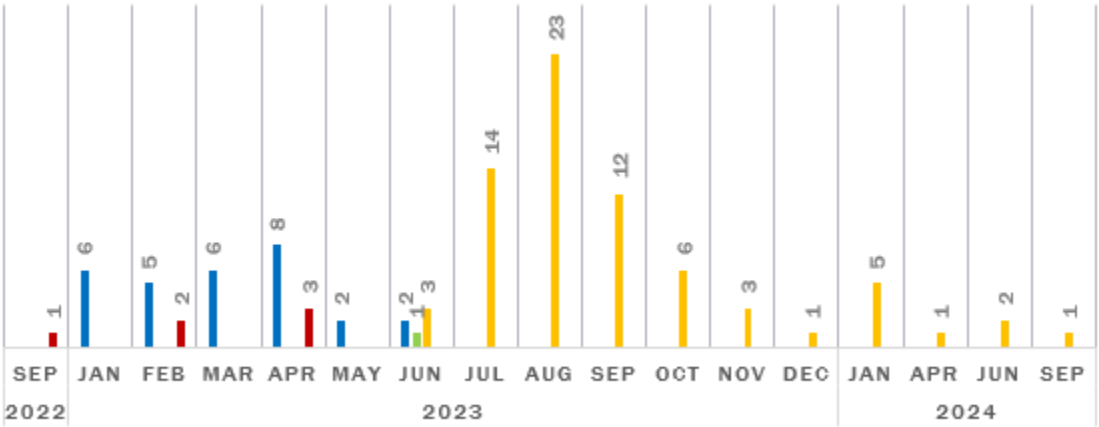
Figure 1.7: Long-term Pivot

The chart above is based on the Long-term pivot table below.

Countdown1		Status				
Years	Start Date	Completed	In Progress	Not Started	Overdue	Grand Total
⊟ 2022	Sep				1	1
2022 Total					**1**	**1**
⊟ 2023	Jan	6				6
	Feb	5			2	7
	Mar	6				6
	Apr	8			3	11
	May	2				2
	Jun	2	1	3		6
	Jul			14		14
	Aug			23		23
	Sep			12		12
	Oct			6		6
	Nov			3		3
	Dec			1		1
2023 Total		**29**	**1**	**62**	**5**	**97**
⊟ 2024	Jan			5		5
	Apr			1		1
	Jun			2		2
	Sep			1		1
2024 Total				**9**		**9**
Grand Total		**29**	**1**	**71**	**6**	**107**
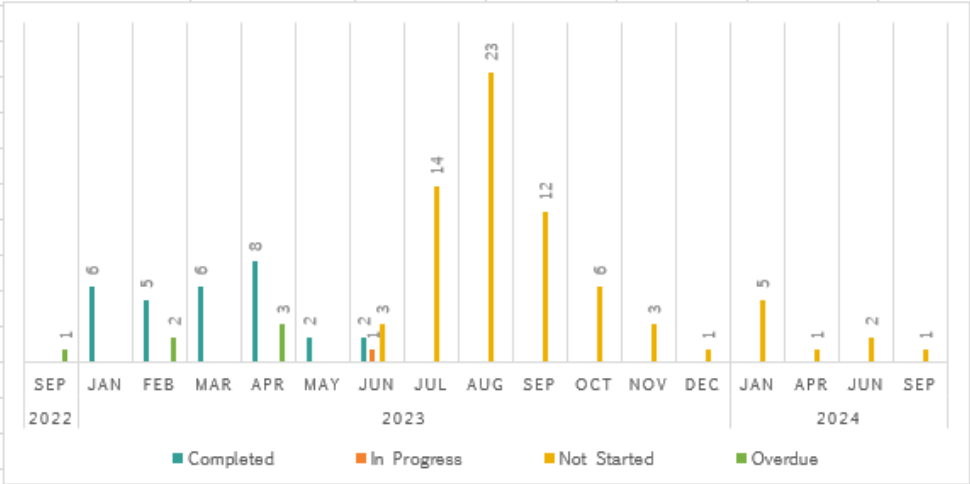
5.0 References

Project files:

https://drive.google.com/drive/folders/1nt1SNRbIOmCyN1KcEeiCMDQELOt_yld5?usp=drive_link