



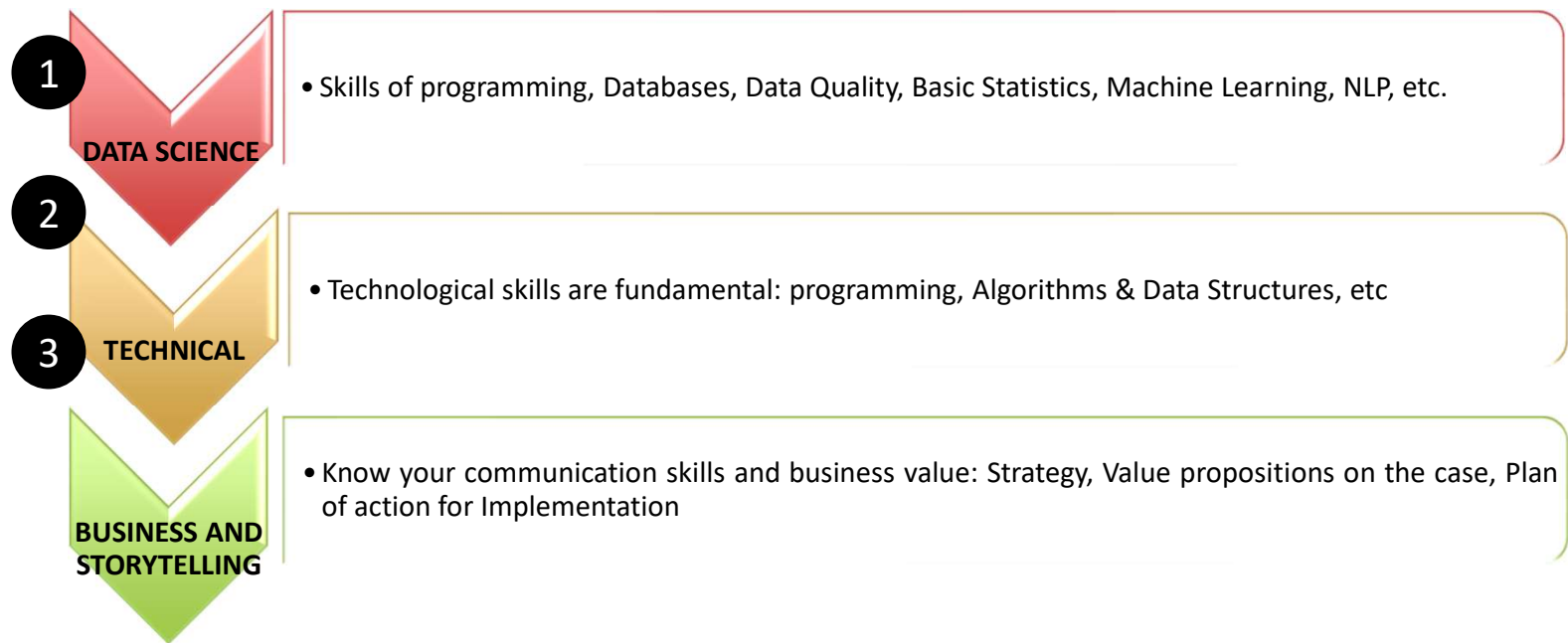
# Data Scientist Technical Test

2019-04-10

# INTRODUCTION

We need to know your technical skills of Big Data Engineer through a practical test. The test does not exclude selection processes but allows us to know your skills.

The test shall assess:



You can sophisticate the response of the case as much as you want

# TECHNICAL TEST

We start with the technical test, consider the following milestones...

## Data Science

- The test will be evaluated by the EA Loc Analytics technical team

## Informed of the case

- Description of the objective and data for the technical test are provided.

## Development of the Test

- Resolution of the case.
- It is recommended to use Open Source tools or free platforms.
- The evaluation of the test will require the delivery of the code (preferably in a **Github repository**)
- You can add any Open Data information you deem appropriate

## Storytelling

- The results will be presented in writing in a repository (preferably Github) and in person to the EA team. Present to us the results obtained in the way you consider convenient. We are open to all options

NOTE: In the case of using tools other than Open Source or Data Sources not free for the case, Electronic Arts (EA) is not responsible for the license, excluding from any responsibility as to the use of the same by the candidate.



# TECHNICAL TEST about FIFA game

**Objective 1:** We need to know what languages are the best for translating the game into many others (this is Localize the game). For that, we have information about the player of one of our games.

The programming exercises must include the installation and execution procedures of the tools to be used and the scripts that will perform the procedures. It can be shared by Github or any repository or in a zip.

## Tasks:

- Loading information: The information provided must be uploaded to some database. It can be structured or unstructured. Include comments on why you chose that type of database. Data repository are:
  - <https://www.kaggle.com/thec03u5/fifa-18-demo-player-dataset> . Here you have information about 70+ attributes, Player and Flag Images, Playing Position Data, Attributes based on actual data of the latest EA's FIFA 18 game, Attributes include on all player style statistics like Dribbling, Aggression, GK Skills etc., Player personal data like Nationality, Photo, Club, Age, Wage, Salary etc.
  - <https://github.com/annexare/Countries/tree/master/data> you have information about languages per country (JSON). Create a new dataset that join the information about countries and country's language and export this dataset to CSV.
- Data Quality: Analyze the quality of the extracted data and transform data if it's necessary.
- EDA: Exploratory data analysis of the FIFA 2018 Player dataset that the main analysis will be the different languages of the players speak.



# TECHNICAL TEST about FIFA game

**Objective 2:** Build a Data Visualization of the data FIFA game by country and languages. Clustering to help us for take the best decisions for Localization.

**Tasks:**

- Faceting, adding visual variables, and summarization
- Profile the different player and teams.

**Objective 3:** Sentiment Analysis of FIFA20

**Tasks:**

- Extract information on social media to obtain information about the new game FIFA 20. We need to detect the sentiment of the localization game for different countries.

## Documentation with the resolution of the case

Sending the file format to evaluate the result, preferable on **Github** platform.

### Code used

The code used for the resolution of the case will be delivered in order to know the process used (methodology) and the degree of knowledge of the tool used for its resolution.

A **Storytelling** should also be provided explaining how the case (methodology) and conclusions of the result obtained have been made. The format of this point is free. The resolution of the case will then be explained in the personal technical interview.

**Test run time:** 1 week is sufficient.



**Good Luck!!!**

