Recruitment Practical Case:

**Understanding telecom fixed penetration trends in Colombia**

**The following skills will be evaluated with the case:**

* Capacity to perform analysis from multidimensional datasets in time series
* Capacity to get insights from data to understand trends and corresponding business implications
* Capacity to clean, join, filter and manipulate datasets in phyton pandas to answer business questions
* Capacity to understand discriminating features for different groups of clusters
* Capacity to prepare an executive presentation in Power Point to summarize most relevant findings
* Capacity to communicate the result of those findings by presenting the resulting material

**Instructions**

The following case is based on real data extracted from the Colombian government portal. It has been prepared to ease the analysis required during the case. The objective is for you to follow the instructions to answer 6 questions mentioned in the python notebook:

1. How is the evolution of the general penetration of fixed internet in Colombia? Are there specific trends that can be identified from the graph? Is the current penetration high or low? What is the maximum penetration that can be reached if it is measured as instructed in this question?
2. Does the evolution of the penetration differ for big and small cities? Is the gap increasing or decreasing (take a clear position and support it with numbers resulting from your analysis)
3. How does average internet offered speeds vary between big cities and small cities? Is the gap increasing or decreasing (take a clear position and support it with numbers resulting from your analysis)
4. Is the distribution between internet service providers different in big cities and small cities? What are the possible implications of that difference? What are the pros and cons of using a pie chart to present this information?
5. Previous questions have shown that two very different groups of cities can be separated by population (size of the city), Can cities be grouped or clustered based on other indicators in the dataset? What are the characteristics from each cluster?
6. Liberty Latin America currently does not have a B2C (Business-to-Consumer) operation in Colombia. **Based on the data provided in the case study**, please provide a recommendation for the most effective entry strategy and identify specific groups of cities for its implementation.

**Required steps**

1. *You will have one week to complete the assignment. Please make sure to have sufficient time to complete the assignment as the delivery dateline cannot be postponed once the notebook is shared.*
2. Install Anaconda (<https://www.anaconda.com/products/distribution>) or any preferred software to run phyton notebooks in your local pc or in a cloud environment
3. Save the dataset *Fix\_Internet\_Tech\_City\_Period.csv.gz* that has been provided with the case in the same folder that you are running your notebook
4. Follow the instructions, get the suggested visualizations, and answer the questions. The notebook will confirm if the output from each step is correct. Your capacity to get insightful and right answers from datasets will be evaluated (content quality). Please make sure that any conclusion, recommendation, or insight is based on your data results.
5. **Recommendation: The notebook will create a database “fixed\_database” in the same folder in the first steps of the code. In case you want to re-start the notebook from scratch because one of the tables has been damaged, you must stop the notebook kernel and delete the database created in the notebook folder.**
6. Prepare a presentation including one slide per question to present your answers. **Please don’t copy and paste the matplotlib graphs on your PPT**, download the resulting tables for each question and include them in your presentation with any modification that supports your argument, possibly including a change in the type of plot to visualize the results. **Each answer to a question shall be limited to a maximum of two slides.** Your presentation will be evaluated based on your capacity to present insightful information in a clean, concise and clear way following best industry practices such as the [Pyramid Principle](https://www.theanalystacademy.com/consulting-slide-structure/). We will also evaluate your capacity to properly setup slide layouts and select a color palette that looks professional and helps to deliver the message.
7. Present your material to the evaluator in a 30-minute session.