**GROUP 1 Challenge 2**

**Cleaning the Artificial Dataset**

**Week 2-3**

Congratulations! You have completed the first Challenge of Creating the Artificial Dataset for the SDS-Fellowship! All your managers are very pleased about your effort!

Now you have to use all your knowledge and innovation in order to “Clean” the Dataset. You are not provided with more information, but they expect from you to fulfill their expectations.

Moreover, they advise you to use **Jupyter Notebook** to present your work, with related comments for every decision you will make.

**At the end of the challenge you will be able to:**

1. Use various methods to clean the Dataset.
2. Clearly present your work to the other Data Scientists.
3. Choose the preferred programming language.
4. Learn how to solve new enquiries online or by collaborating with other students.
5. Learn to think out of the box by creating your own path throughout the challenge.
6. Learn how to work into specific deadlines.
7. Use Jupyter Notebook for Data Science projects. (optional)

This is the **second step of GROUP 1** of the series of **SDS Fellowship Challenges** (See the full SDS Fellowship Plan [here](https://docs.google.com/document/d/1wzGMoJlJDMtpJVqb05u4Ea-ADFfk1RgNZVrer2aJCeU/edit?usp=sharing)). After completing this task the students who collects the most "likes" from fellow students will be announced as the **Junior Data Scientist of the Month**!

**Dates of Participation:**

**Working on the data cleaning: Monday, October 21st - Sunday, October 27th.**

**Deadlines:**

1. **Mon, October 21st - Sun, October 27th:** work on the challenge on your own or in small groups of students. Feel free to discuss the challenge but **please don’t post your results earlier than October 27th!**
2. **Mon, October 27th:** present your results to the channel.
3. **Mon, October 28th - Wed, October 30th:** reviewing and voting for the work of fellow students
4. **Thur, October 3st**: announcing the **Junior Data Scientist of the Month**.

**The dataset to use:**

Most of the members voted for [@Alex](https://sdscommunity.slack.com/team/ULZD5ESKU)’s dataset to use in the upcoming challenges, so that's what we go further with.

Here is a link to his GitHub: <https://github.com/korotulea/SDS-FELLOWSHIP-CHALLENGE>

Or here is his csv file: <https://github.com/korotulea/SDS-FELLOWSHIP-CHALLENGE/blob/master/financial_artificial_dataset_AK.csv>

**Are you ready? :rocket::rocket::rocket:**

**Let’s continue the journey!**

**Here is the Slack channel for the Challenge discussions:** <https://sdscommunity.slack.com/messages/CNVFRNLR3>

**Send a private message to your Community Manager if you have any questions or difficulties:**

<https://sdscommunity.slack.com/messages/DL5GEQP4N>