Bank credit scoring

Scoring – client rate system that uses statistics methods. Usually, it is algorithm that takes client’s data. In result we get answer – should we credit them or should not.

**Data:**

|  |  |
| --- | --- |
| **Column** | **Description** |
| client\_id |  |
| age |  |
| sex |  |
| married |  |
| salary | Official confirmed salary |
| successfully\_credit\_completed | Number of completed credits |
| credit\_completed\_amount | Total sum of completed credits |
| active\_credits | Number of active credits |
| active\_credits\_amount | Total sum of active credits |
| credit\_amount | Credit Sum (loan) |
| is\_credit\_closed |  |

Program must rate clients

Requirements:

1) [Python 3.8+](https://www.python.org/downloads/)

2) [PyCharm](https://www.jetbrains.com/ru-ru/pycharm/) / Jupyter

3) (for local PC):

pip install pyarrow==5.0.0 pandas==1.3.3 scikit-learn==1.0 pyspark==3.1.2

4) configured JDK not above 11!   
tutorial [here](https://skillbox.ru/media/base/kak_ustanovit_jdk_i_sredu_razrabotki_intellij_idea/).