|  |
| --- |
| EPAM Systems, RD Dep. |
| BI.Talend Data Integration Labwork 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| REVISION HISTORY | | | | | |
| Ver. | Description of Change | Author | Date | Approved | |
| Name | Effective Date |
| 1 | Initial version | Sergey Moukavoztchik | 12/01/2018 |  |  |
|  |  |  |  |  |  |

Contents

[1. Preparation 3](#_Toc503489527)

[2. Tasks 3](#_Toc503489528)

# Preparation

1. Download and unarchive Talend Open Studio for Data Integration.
2. Create project
3. Created jobs and other repository elements should be Exported and pushed to separate repo at github.
4. Link to the repo need to be added to google sheet https://docs.google.com/spreadsheets/d/1PaA-xGj1087e0Uzd5vstPeCj8e198EyUO2j-3s\_9DK4/edit?usp=sharing
5. Optionally you can add text document with a description of your jobs.

# Tasks

Three files provided customers.csv, products.csv, payments.csv.

* Create metadata for each of them.
* Create a file where listed top 20 buyers (by amount spent).
* Create a file where listed top 20 companies by their revenue.
* Create file with list of payment records which was created for products with missing price, this records should not be considered in previous statistics