**Data Analysis**

**Task 3**

Worldbank\_upd.xlsx file contains World Bank data by countries for 2018 and 2019.

Source: <https://databank.worldbank.org/source/world-development-indicators>

**Description of variables**

* CountryName - Country Name
* CountryCode - Country Code
* high\_income - High income countries (Takes 2 values: 1 - belongs to the high income group of countries, 0 - no)
* income - Income (Takes 3 values: high income, middle income, low income)
* development - Developed or developing country (Takes 2 values: developing, developed)

**Variables with data for 2018:**

* Agriculturalrawmaterialsexpor - Agricultural raw materials exports (% of merchandise exports)
* ArmsexportsSIPRItrendindica - Arms exports (SIPRI trend indicator values)
* Commercialserviceexportscurr - Commercial service exports (current US dollars)
* Communicationscomputeretc - Communications, computer, etc. (% of service exports, BoP)
* Exportsofgoodsandservices - Exports of goods and services (% of GDP)
* Exportsofgoodsandservicesa - Exports of goods and services (annual percent growth)
* ExportsofgoodsandservicesB - Exports of goods and services (BoP, current US dollars)
* Exportsofgoodsandservicesc - Exports of goods and services (constant 2010 US dollars)
* Foodexportsofmerchandisee - Food exports (% of merchandise exports)
* GoodsexportsBoPcurrentUS - Goods exports (BoP, current US dollars)
* Fuelexportsofmerchandisee - Fuel exports (% of merchandise exports)
* Hightechnologyexportsofma - High-technology exports (% of manufactured exports)
* Hightechnologyexportscurrent - High-technology exports (current US dollars)
* TaxesonexportscurrentLCU - Taxes on exports (current LCU)
* Transportservicesofservice - Transport services (% of service exports, BoP)
* Travelservicesofcommercial - Travel services (% of commercial service exports)
* Travelservicesofserviceex - Travel services (% of service exports, BoP)

**Variables with data for 2019**

* Travelservicesofcommercial2019 - Travel services (% of commercial service exports)
* Agriculturalrawmaterialsexpo2019 - Agricultural raw materials exports (% of merchandise exports)
* Communicationscomputeretc2019 - Communications, computer, etc. (% of service imports, BoP)
* Transportservicesofservice2019 - Transport services (% of service exports, BoP)

**Please consider the following tasks.**

Analyze two variables:

* Communications, computer, etc. (% of service imports, BoP);
* Transport services (% of service exports, BoP).

Select appropriate tests to check if there is a difference in values of each variable in the following groups:

* High income and other countries;
* Countries related to different income groups (high, middle, low);
* Developing and developed countries.

Compare the values of each variable for different years (2018 and 2019).

Explain the selection of each statistical test. Formulate the hypotheses. Interpret the results of analysis.

Please send the Jupyter Notebook with comments and answers to amelikyan@hse.ru from your **personal** e-mail. The task can be done individually or in a group of 2-3 students. If the task is done in a group, then one member of the group can send the task indicating the names of the group members in the subject of the email.