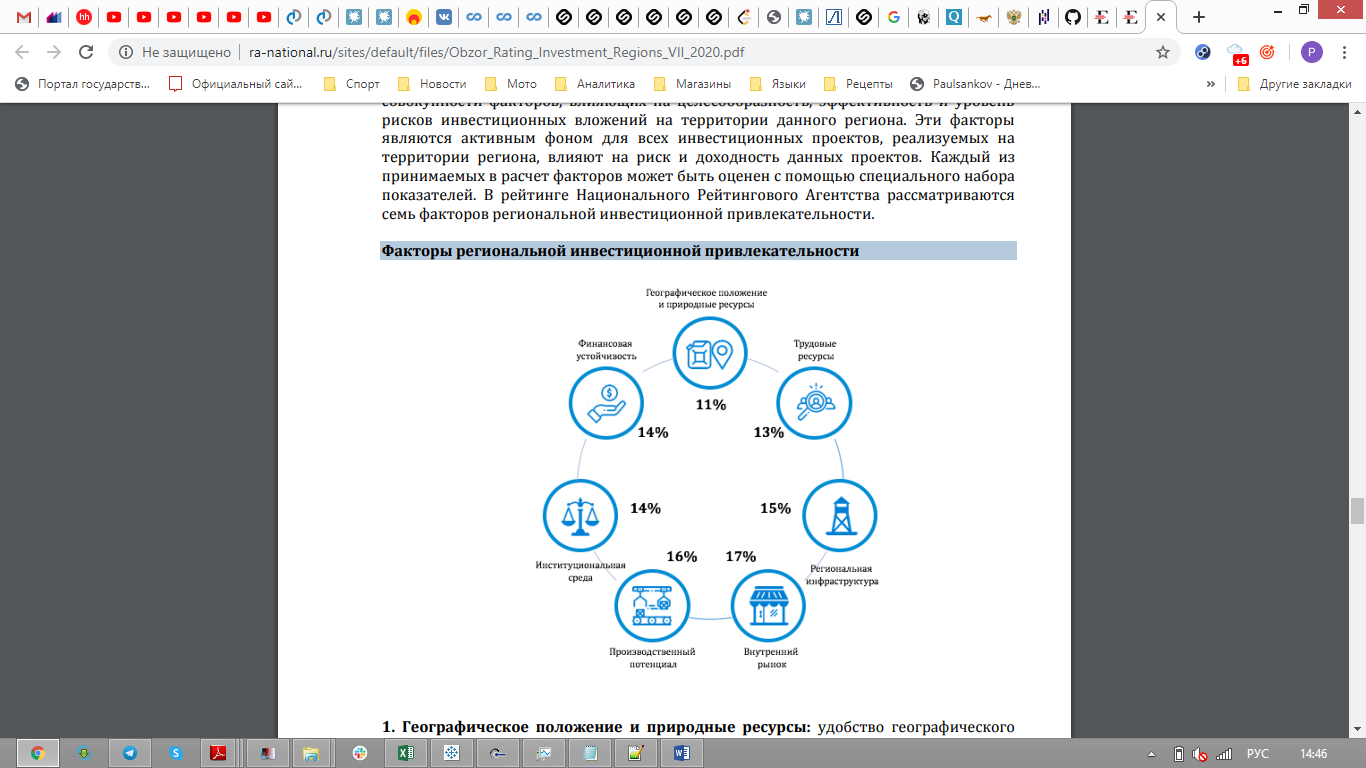
**ОЦЕНКА ИНВЕСТИЦИОННОЙ ПРИВЛЕКАТЕЛЬНОСТИ РЕГИОНА**

# **Концепция:**

Параметры оценки инвестиционной привлекательности регионов (ИПР) выбирал под влиянием указанных статей:

<https://cyberleninka.ru/article/n/kak-otsenit-investitsionnuyu-privlekatelnost-regiona-ekspertnyy-i-statisticheskiy-metody-otsenki-investitsionnoy-privlekatelnosti/viewer>

<https://cyberleninka.ru/article/n/ekspress-analiz-investitsionnoy-privlekatelnosti-regionov/viewer>



Так как коеффициент базовых усилий, инвест.климата и риска имеют под капотом также достаточно сложный и спорный метод оценки, относил к возможному признаковому описанию параметры, так или иначе отражающиеся на аспектах с картинки (не претендую после прочтения 2 статей на авторитет и объективную оценку в этой области) и **выбрал два пути решения** поставленной задачи:

- среди выбранных параметров для каждого региона оценивать инвестиционную привлекательность по опыту прошлых лет регрессионным методом;

- для признаков, измеряемых хотя бы раз в квартал, строить временной ряд до интересующей даты (ИД), для признаков с годовыми значениями брать среднее двух предыдущих лет от ИД.

## **Признаковое описание регионов**

Поскольку таблицы весьма разнородны как <https://tverstat.gks.ru/storage/mediabank/inv3-2019(1).htm> или <https://clck.ru/MNj9x> или попросту не работают <http://db.novosibstat.ru/dbs/sibdb/dbinet.cgi?pl=2224055> наиболее достоверным и основным источником данных оказалась БИР-Аналитик (реально годный и достаточный инструмент).

Данные <https://data.gov.ru/> и <https://fedstat.ru/>, как и региональных порталов (по крайней мере выбранных 3 регионов), чаще всего даже не доходят до среза 2018 года.

Для разнообразия включил Индексы тарифов на грузовые перевозки с showdata.gks.ru

<https://showdata.gks.ru/report/274438/?filter_1_0=2016-01-01+00%3A00%3A00%7C-56%2C2016-02-01+00%3A00%3A00%7C-56%2C2016-03-01+00%3A00%3A00%7C-56%2C2016-04-01+00%3A00%3A00%7C-56%2C2016-05-01+00%3A00%3A00%7C-56%2C2016-06-01+00%3A00%3A00%7C-56%2C2016-07-01+00%3A00%3A00%7C-56%2C2016-08-01+00%3A00%3A00%7C-56%2C2016-09-01+00%3A00%3A00%7C-56%2C2016-10-01+00%3A00%3A00%7C-56%2C2016-11-01+00%3A00%3A00%7C-56%2C2016-12-01+00%3A00%3A00%7C-56%2C2017-01-01+00%3A00%3A00%7C-56%2C2017-02-01+00%3A00%3A00%7C-56%2C2017-03-01+00%3A00%3A00%7C-56%2C2017-04-01+00%3A00%3A00%7C-56%2C2017-05-01+00%3A00%3A00%7C-56%2C2017-06-01+00%3A00%3A00%7C-56%2C2017-07-01+00%3A00%3A00%7C-56%2C2017-08-01+00%3A00%3A00%7C-56%2C2017-09-01+00%3A00%3A00%7C-56%2C2017-10-01+00%3A00%3A00%7C-56%2C2017-11-01+00%3A00%3A00%7C-56%2C2017-12-01+00%3A00%3A00%7C-56%2C2018-01-01+00%3A00%3A00%7C-56%2C2018-02-01+00%3A00%3A00%7C-56%2C2018-03-01+00%3A00%3A00%7C-56%2C2018-04-01+00%3A00%3A00%7C-56%2C2018-05-01+00%3A00%3A00%7C-56%2C2018-06-01+00%3A00%3A00%7C-56%2C2018-07-01+00%3A00%3A00%7C-56%2C2018-08-01+00%3A00%3A00%7C-56%2C2018-09-01+00%3A00%3A00%7C-56%2C2018-10-01+00%3A00%3A00%7C-56%2C2018-11-01+00%3A00%3A00%7C-56%2C2018-12-01+00%3A00%3A00%7C-56%2C2019-01-01+00%3A00%3A00%7C-56%2C2019-02-01+00%3A00%3A00%7C-56%2C2019-03-01+00%3A00%3A00%7C-56%2C2019-04-01+00%3A00%3A00%7C-56%2C2019-05-01+00%3A00%3A00%7C-56%2C2019-06-01+00%3A00%3A00%7C-56%2C2019-07-01+00%3A00%3A00%7C-56%2C2019-08-01+00%3A00%3A00%7C-56%2C2019-09-01+00%3A00%3A00%7C-56%2C2019-10-01+00%3A00%3A00%7C-56%2C2019-11-01+00%3A00%3A00%7C-56%2C2019-12-01+00%3A00%3A00%7C-56%2C2020-01-01+00%3A00%3A00%7C-56&filter_2_0=120418&filter_3_0=188254&filter_4_0=13142%2C13268%2C13238&rp_submit=t> и расходы на реализацию мер социальной поддержки граждан

<https://showdata.gks.ru/report/274240/?&filter_1_0=2016-06-01+00%3A00%3A00%7C-56%2C2016-12-01+00%3A00%3A00%7C-56%2C2017-06-01+00%3A00%3A00%7C-56%2C2017-12-01+00%3A00%3A00%7C-56%2C2018-06-01+00%3A00%3A00%7C-56%2C2018-12-01+00%3A00%3A00%7C-56%2C2019-06-01+00%3A00%3A00%7C-56&filter_2_0=224194&filter_3_0=222267&filter_4_0=130519&filter_5_0=13142%2C13268%2C13238&rp_submit=t>

Насчет парсинга данных с этого сайта были преодолимые трудности, требующие несоизмеримой с эффектом траты времени, скачал в формате эксель.

## **Целевую переменную** – ИПР – выбрал по годам место региона в рейтинге:

<http://www.ra-national.ru/sites/default/files/Obzor_Rating_Investment_Regions_VII_2020.pdf>

<http://www.ra-national.ru/sites/default/files/analitic_article/IPR-6-06112018.pdf>

<http://www.ra-national.ru/sites/default/files/analitic_article/%D0%98%D0%BD%D0%B2%D0%B5%D1%81%D1%82%D1%80%D0%B5%D0%B3%D0%B8%D0%BE%D0%BD%D1%8B%202017.pdf>

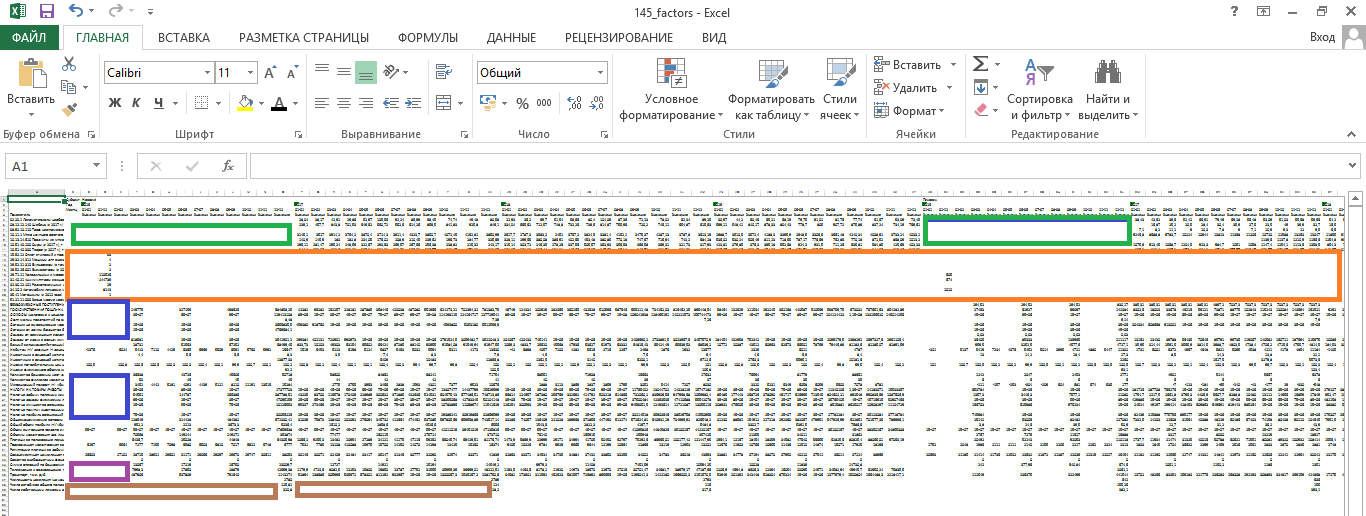
<http://www.ra-national.ru/sites/default/files/analitic_article/%D0%9D%D0%A0%D0%90_%D0%A0%D0%B5%D0%B9%D1%82%D0%B8%D0%BD%D0%B3%20%D0%B8%D0%BD%D0%B2%D0%B5%D1%81%D1%82%D0%B8%D1%86%D0%B8%D0%BE%D0%BD%D0%BD%D0%BE%D0%B9%20%D0%BF%D1%80%D0%B8%D0%B2%D0%BB%D0%B5%D0%BA%D0%B0%D1%82%D0%B5%D0%BB%D1%8C%D0%BD%D0%BE%D1%81%D1%82%D0%B8%20%D1%80%D0%B5%D0%B3%D0%B8%D0%BE%D0%BD%D0%BE%D0%B2%20%D0%A0%D0%A4%202016_%D0%98%D0%A2%D0%9E%D0%93_0.pdf>

Поскольку мы не знаем исходя из этого рейтинга насколько большая разница в показателях любых двух позиций, то это еще одно грубое допущение в расчете. Поэтому нашел более адекватную величину, хоть и устаревшую, но все же позволяющую оценить разброс в рейтинге и получить более-менее привязанный к цифрам показатель:

<https://www.raexpert.ru/rankings/region_climat/2017>

# **Подготовка данных**

Еще на этапе изучения сырых данных из БИР видно много пропусков в признаках:



- часть из них – заполнялись по кварталам или раз в год, некоторые имеют историю с 2017 года;

- другую часть, имеющую одно-два значения за весь период придется дропнуть, особенно просто, пока в экселе. Впоследствии можно выгружать выборку без них.

Также отмечу, что среди признаков имелись ненулевые только для одного региона – это наводит на очень важную мысль: нельзя обучить регрессионную модель на признаках и целевой переменной одного региона и применить к другому, они различны и потому для каждого региона (при применении регрессионной модели предсказания ИПР) перечень признаков был бы свой. Если бы была известна универсальная формула по которой рассчитывается ИПР, возможно она бы давала более точный показатель по сравнению с ra-national.ru. Но не факт, что там учитываются все особенности конкретного региона. К тому же цель нашего исследования другая, нежели стандартизация алгоритма оценки ИПР.

Продолжение в jupiter notebook.