EDA and links.

*Цель: собрать и проанализировать данные, имеющие возможную корреляцию с уровнем воды в реках и озерах Восточно Казахстанской Области.*

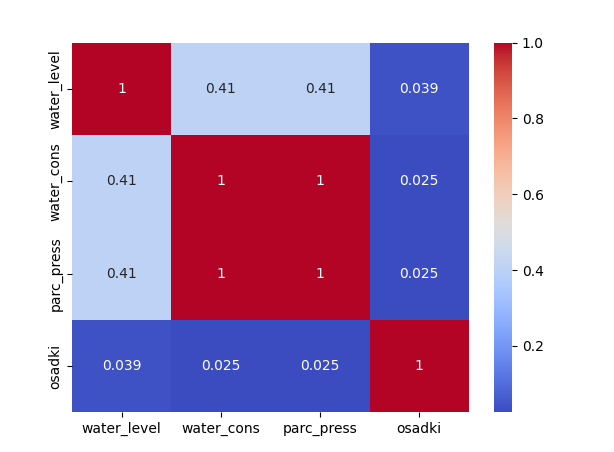
Во время разведаточного анализа была рассмотрены корреляции между следующими входными данными: Уровень воды,

расходы воды, атмосферные осадки, атмосферное давление, температура (максимальная, средняя, минимальная), толщина льда и снега (данные были взяты из ЕДК Казгидромет), дефицит насыщения, облачность, средняя скорость ветра, а также цикличность (сезонность) изменений.

Метереологические данные для отдельных районов были взяты с метеорологической и гидрологической БД Казгидромет. (<https://www.kazhydromet.kz/ru/interactive_cards>)

Однако во время изучения корреляции было выявлено, что несмотря на предпологаемую очевидность существования линейной связи и корреляции между перечисленными данными и уровнем воды, все признаки, кроме **уровня воды, количества атмосферных осадков, парциального давления, расходов воды, а также цикличности**, имели отрицательную либо крайне низкую (меньше 0.01) корреляцию, и не имели большого влияния на обучение моделей.

Корреляция между наиболее выбранными признаками одной из рек (р.Абылайкит – с.Самсоновка) отражена на следующей тепловой карте *рис.1* (подробнее можно ознакомиться в cor-matrix.py):



*рис. 1*