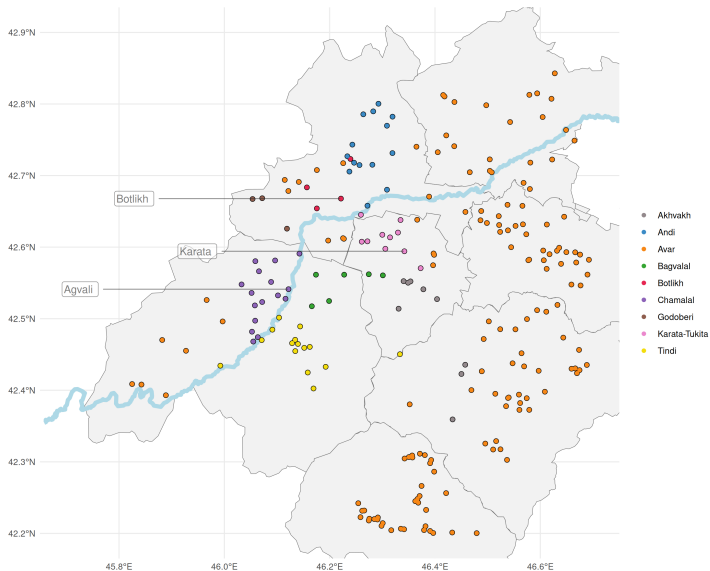


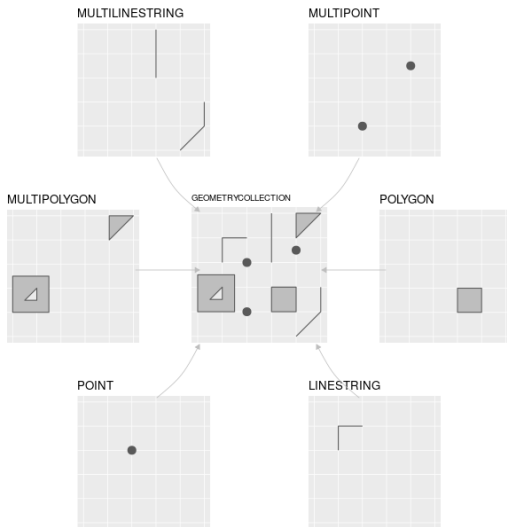
Анализ данных может включать

- сбор данных
- очистку данных и их предобработку
- визуализацию данных
- моделирование данных
- дескриптивный анализ
- предиктивный анализ
- машинное обучение
- ...

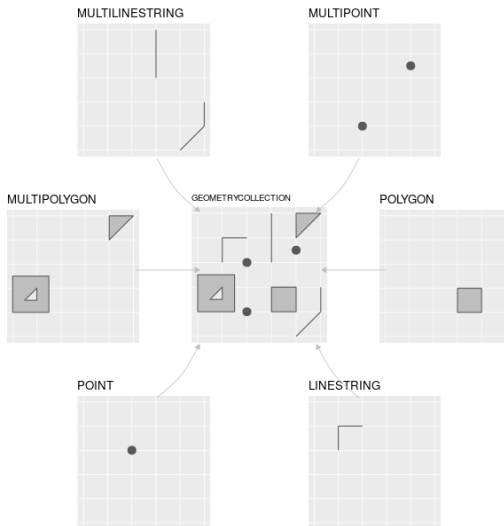
Какие пространственные примитивы можно здесь найти?



Чего, как Вам кажется, здесь не хватает?



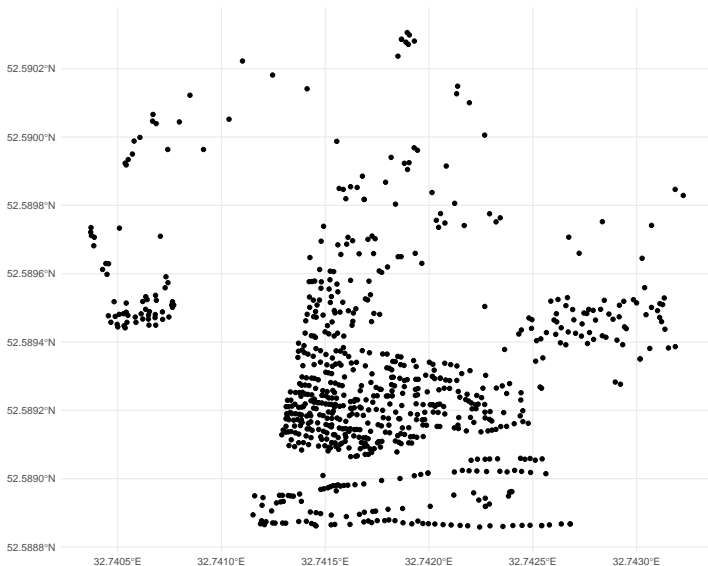
Чего, как Вам кажется, здесь не хватает?



Мне не хватает объема (т. е. учета высотности).

Растровые данные

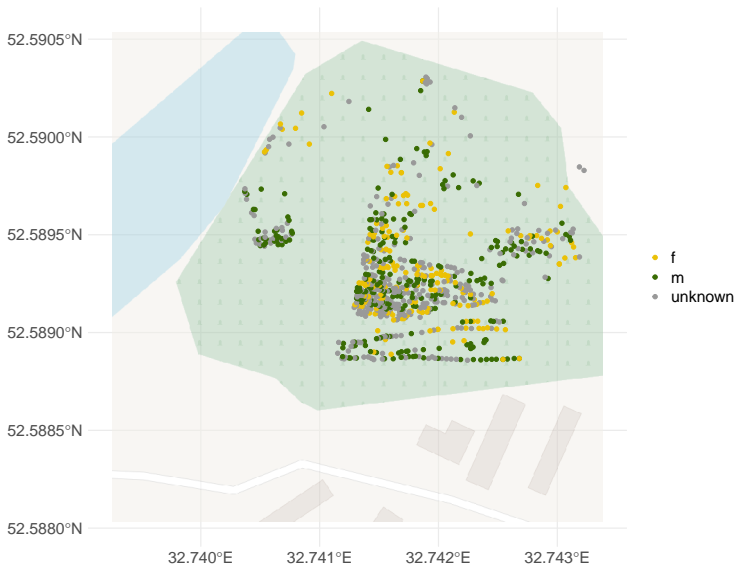
Кладбище Стародуб (данные полевого архива SFIRA)



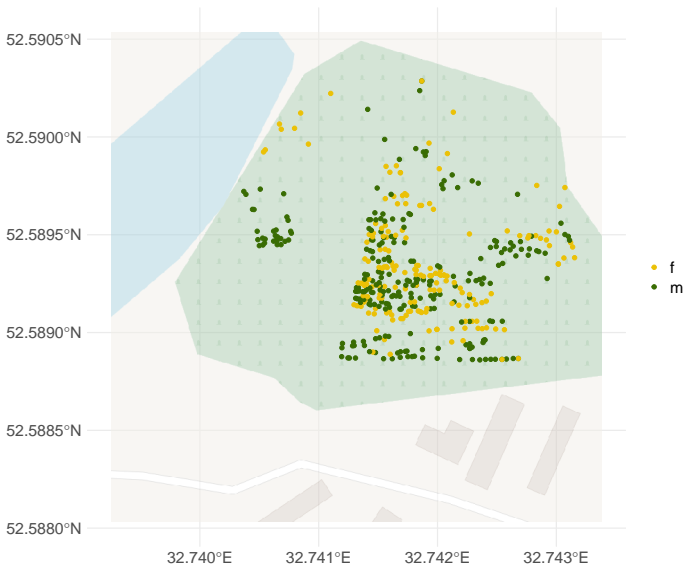
Кладбище Стародуб (данные полевого архива SFIRA)



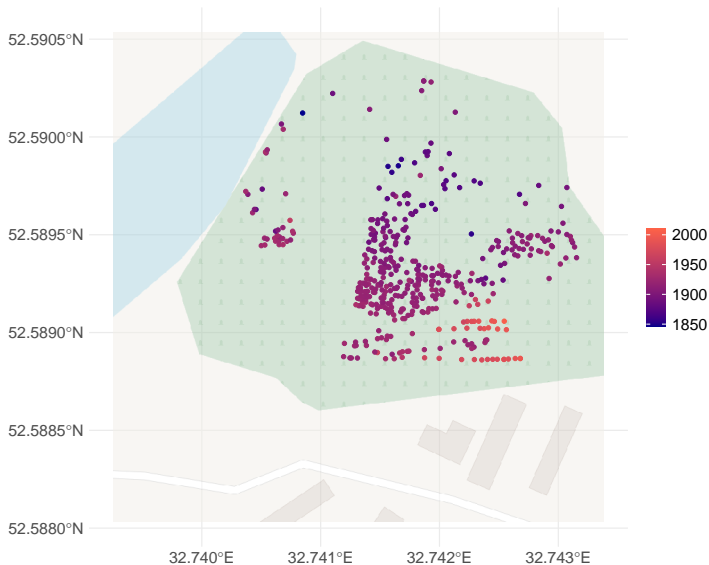
Кладбище Стародуб (данные полевого архива SFIRA)



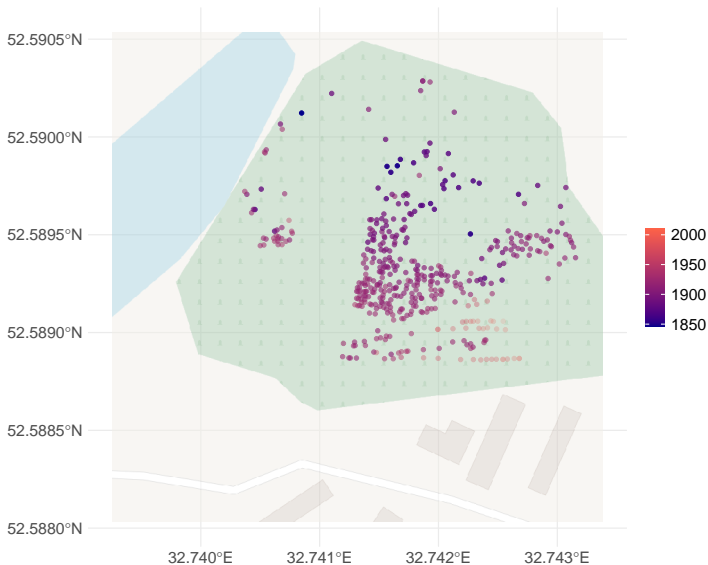
Кладбище Стародуб (данные полевого архива SFIRA)



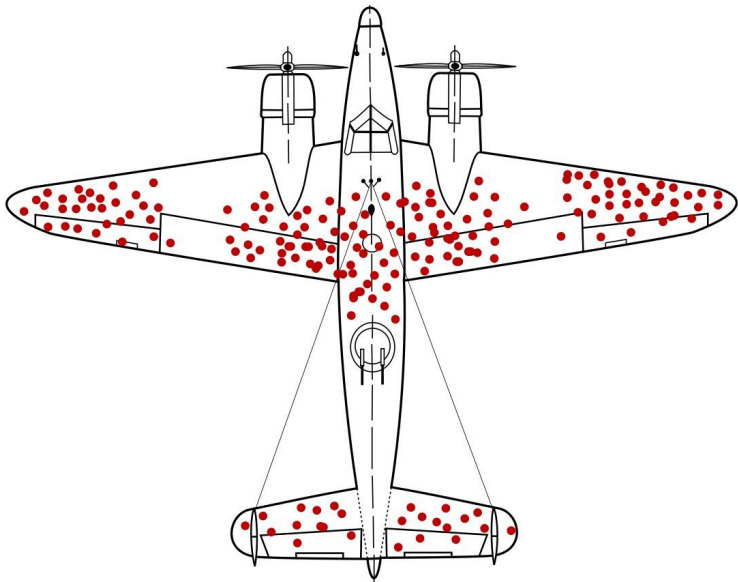
Кладбище Стародуб (данные полевого архива SFIRA)



Кладбище Стародуб (данные полевого архива SFIRA)



Ошибка выжившего: Абрахам Вальд



Картографическая проекция

Любое отображение некоторого небесного тела на плоскость называют картографической проекцией.

Если расстояния в ваших данных небольшие (особенно, если координаты близки к экватору), широту и долготу можно без страха использовать как оси в декартовой системе координат (она же — проекция Меркатора). Однако при работе с данными масштаба страны/континента/планеты такой подход будет накапливать ошибку из-за искажений одного из следующих типов:

- искажения длин;
- искажения углов;
- искажения площадей;
- искажения форм.

Картографическая проекция

Проекция Меркатора очень сильно искажает площади:



исходный



с сохранением площадей

источник — Википедия

Картографическая проекция

- веб-приложение, помогающее выбрать подходящую проекцию
- веб-приложение, которое показывает как изменяются объекты при преобразовании с сферы на одну из четырех проекций (Меркатора, цилиндрическую, Робинсона, Моллвейде)
- Здесь содержится список всех возможных проекций

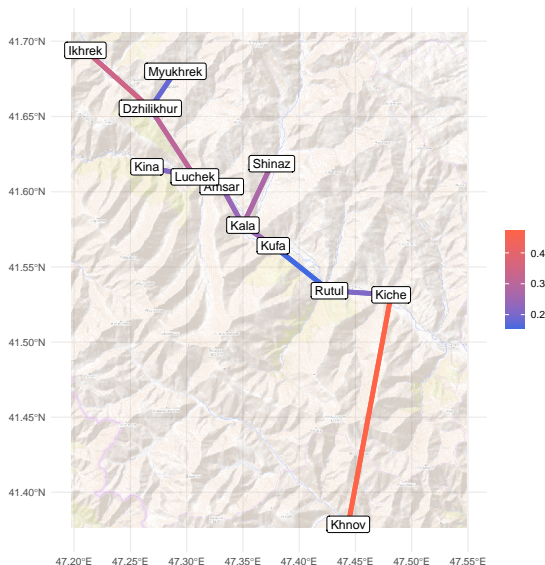
Моделирование пространственных отношений

Моделирование пространственных отношений позволяет отвечать на вопросы:

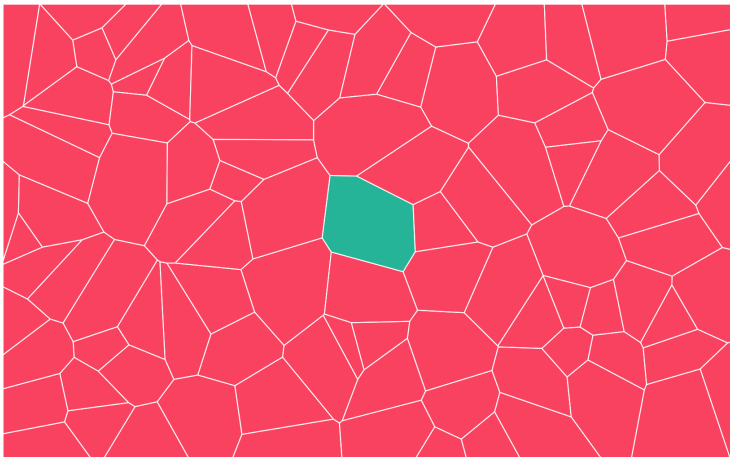
- Существует ли какая-то группировка значений исследуемой переменной в пространстве?
- Правда ли, что сходные значения имеют тенденцию находиться рядом?
- Можно ли выделить какие-то регионы концентрации каких-то из значений?

Однако для ответа на все эти вопросы мы прежде всего должны построить граф соседства.

Языковое сходство рутульских идиомов

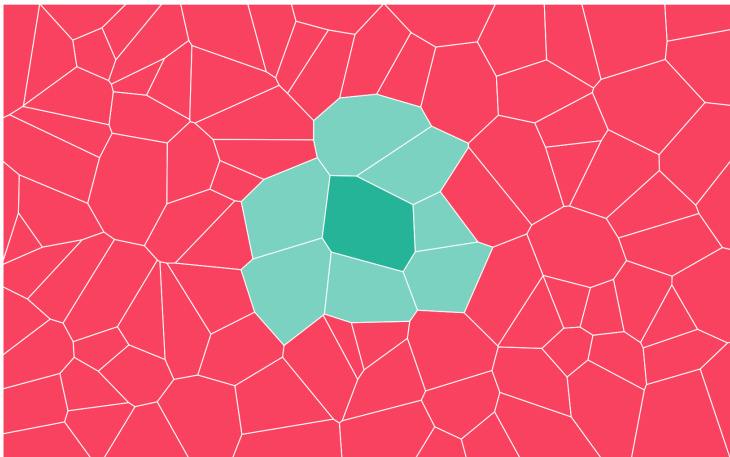


Как определить соседей?



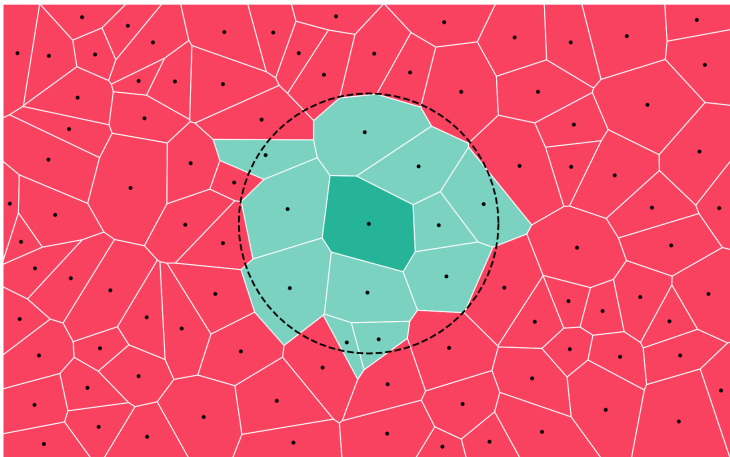
Из курса М. Фляйшманна “Spatial Data Science for Social Geography”

Как определить соседей?



Из курса М. Фляйшманна “Spatial Data Science for Social Geography”

Как определить соседей?



Из курса М. Фляйшманна “Spatial Data Science for Social Geography”

Пространственная автокорреляция

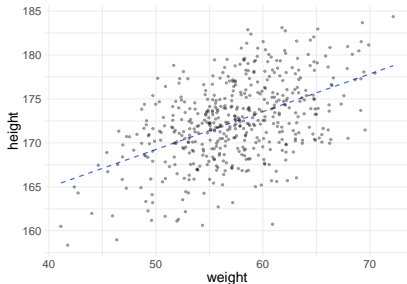
Для начало нужно разобраться с понятием *корреляции*.

Корреляцией называют некоторую связь числовых переменных, если при изменении значения одной переменной систематически изменяется другая переменная.

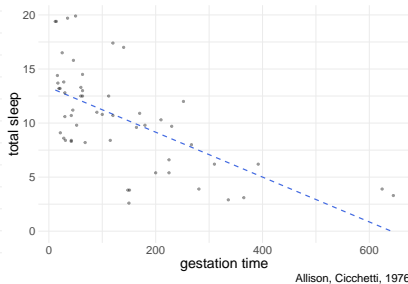
Пространственная автокорреляция

Для начало нужно разобраться с понятием *корреляции*.

Корреляцией называют некоторую связь числовых переменных, если при изменении значения одной переменной систематически изменяется другая переменная.



положительная корреляция



Allison, Cicchetti, 1976

отрицательная корреляция

Пространственная автокорреляция

Степень в какой сходные значения находятся рядом.

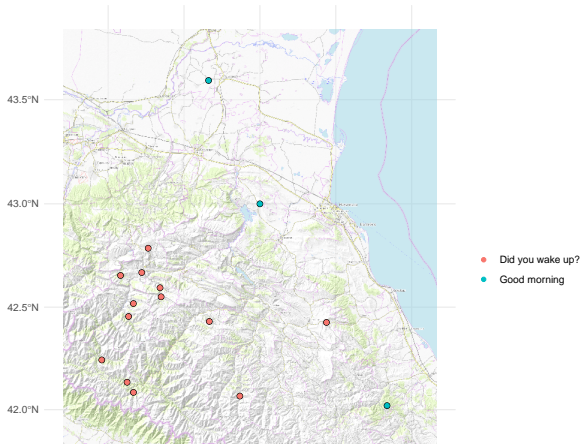
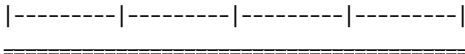
- положительная автокорреляция: похожие значения находятся рядом
- отрицательная автокорреляция: похожие значения находятся далеко друг от друга

Пространственная автокорреляция

Степень в какой сходные значения находятся рядом.

- положительная автокорреляция: похожие значения находятся рядом
- отрицательная автокорреляция: похожие значения находятся далеко друг от друга
- глобальная: имеют ли значения тенденцию оказываться рядом с другими похожими/непохожими значениями;
- локальная: существует ли некоторая специфический фрагмент пространства, где наблюдается необычная концентрация похожими/непохожих значений.

Значение Moran I:



Мне хочется выразить благодарность
Евгению Николаевичу Матерову за его блог и телеграм-канал “Наука
и данные” (<https://t.me/naukaidannye>), которые значительно
упростили написание этой лекции, в частности за ссылку на курс
Мартина Фляйшманна “Spatial Data Science for Social Geography”.



Временные данные