

**Количественный анализ
массового
эпиграфического
материала:
методика применения и интерпретации**

Станислава Хижнякова

Содержание доклада

- Зачем это нужно?
- Проблемы количественного анализа эпиграфики
- Создание датасета
- Подготовка материала
- Обработка материала
- Презентация результатов
- Интерпретация результатов

Зачем это нужно?

- Изучение эпиграфической традиции
- Сравнение традиций друг с другом



- Эволюция демографических тенденций
- Миграционные модели
- Экономическое развитие
- Культурное развитие
- Изменения социально-политической организации

Проблемы

- А какая у нас выборка?
- Латинские и греческие надписи: разрозненные корпуса
 - А у нас...
 - => делать корпуса

Проблемы

- Сложность согласования

=> подразделы данных

- Подготовка материала

=> regular expressions

Проблемы

- Недостаточность данных

- Геометки

=> привязка к памятнику

- Прорисовки

=> CNN

(но тут одни сплошные проблемы)

Проблемы

- Разные типы данных

=> разные библиотеки

Создание датасета

Useful

- Информация о тексте надписи
- Информация о носителе надписи
- Информация о провенансе

Useless

- Информация о месте хранения
- Библиография

Подготовка материала

- Равномерное распределение данных

Graffito

Honorific

Honorific



Graffito

Honorific

Unclassified

Unclassified

Honorific

Подготовка материала

- Подразделы данных

Syria	Palmyra	Valley of Tombs	Tower 63
Syria	Palmyra	N-W necropolis	Tower 155
Syria	Palmyra	Valley of Tombs	Tower 51
Syria	Palmyra	Valley of Tombs	Tower 51
Syria	Palmyra	Valley of Tombs	Tower 68

Подготовка материала

- Упрощение данных

terracotta & clay?

- Унификация данных

100AD | AD 100 | 100 | 100 н. э.

Обработка материала

Как считать среднее?

- mean — даты
- median — размер текста
- mode — категориальные переменные

Обработка материала

Как делить на периоды?

- равномерные — для создания полного графика
- исторические — для сравнения друг с другом

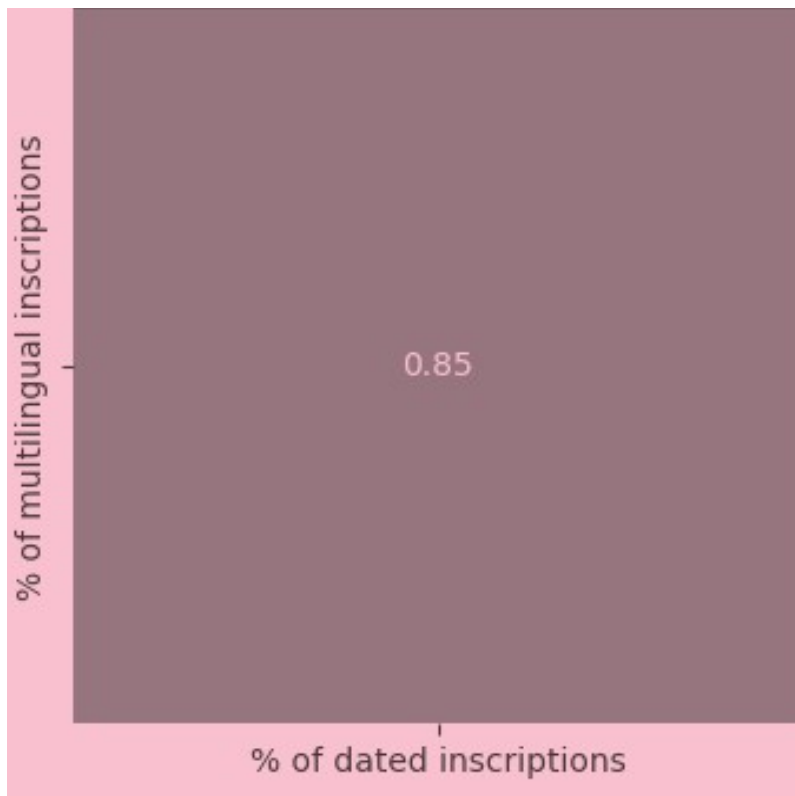
Что делать с выбросами?

- удаление
- интерпретация

Обработка материала

Как искать корреляции?

- категориальные переменные => количественные



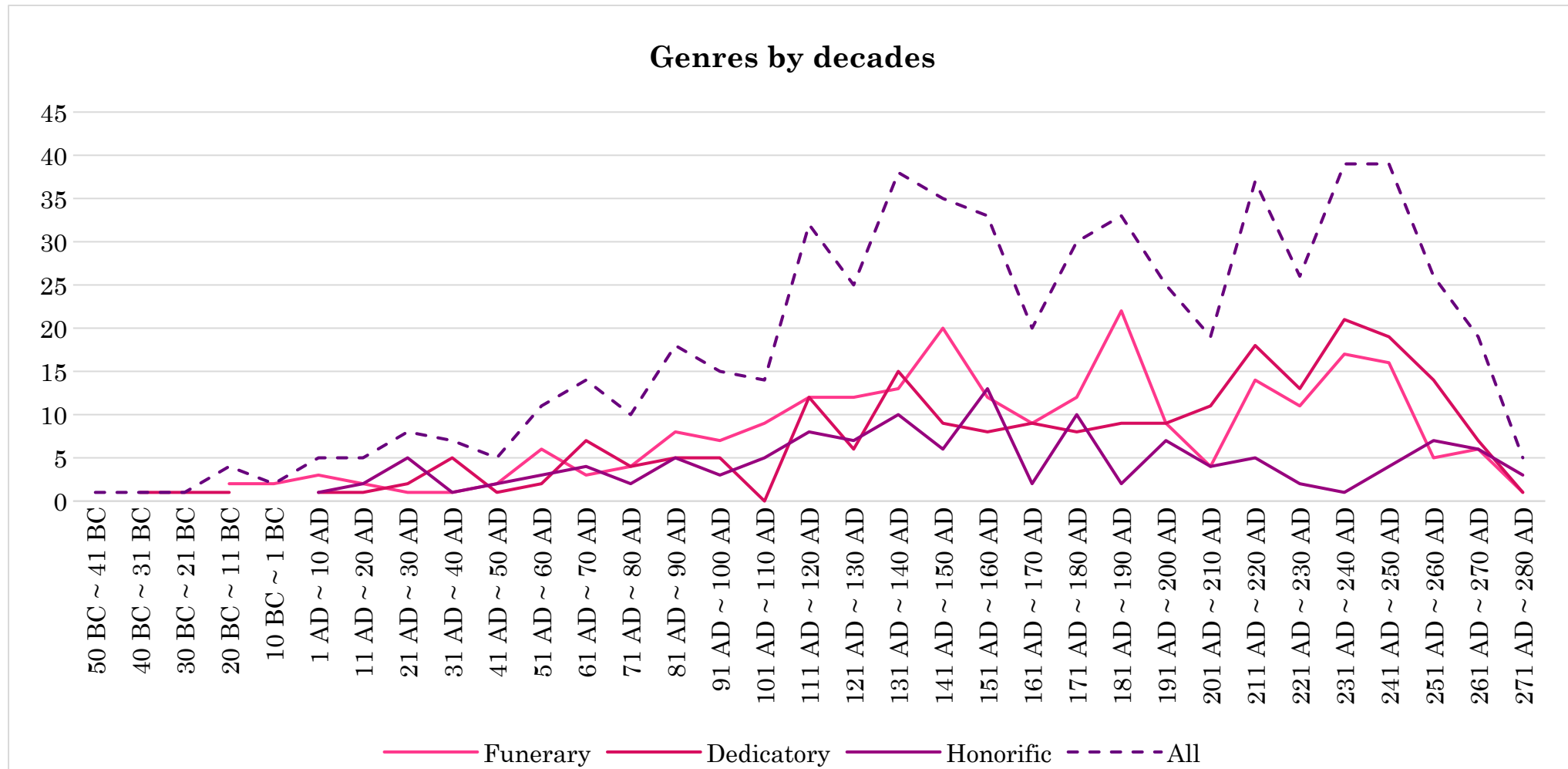
Обработка материала: библиотеки

- pandas – для основной статистической обработки
- geopandas – для работы с геометками
- nltk – для обработки естественного языка
- temprun – для работы с датировками (метод MCS)
- numru – для высокоуровневой математической обработки данных

Презентация результатов: библиотеки

- seaborn — для создания графиков
- matplotlib — для создания красивых графиков

Презентация результатов: графики

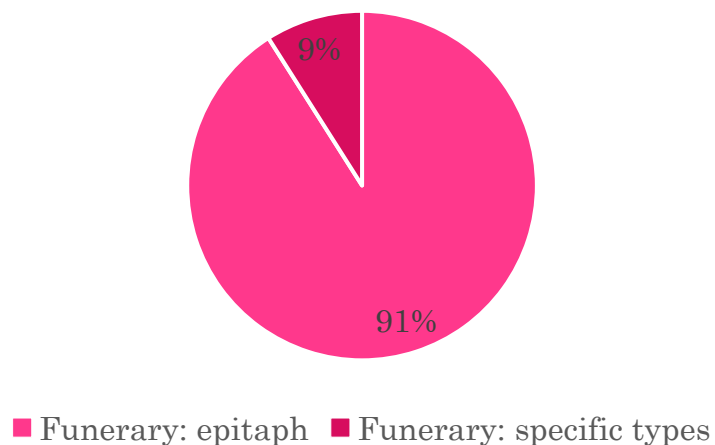


Презентация результатов: графики

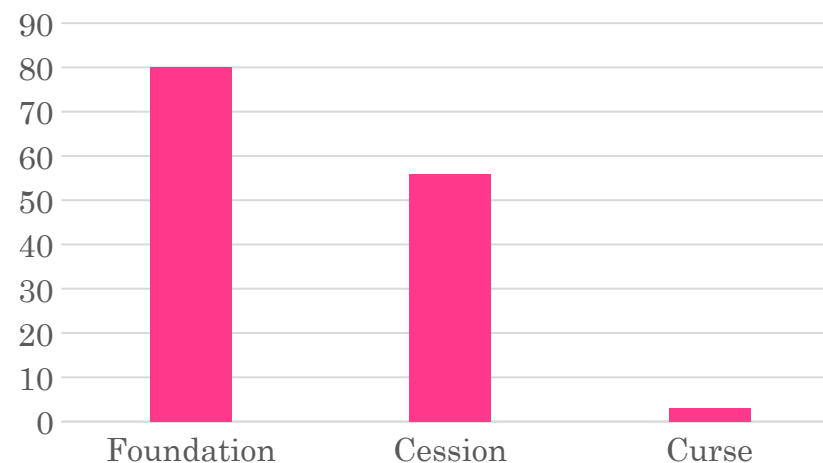
- Не видны экстремумы

=> разбить данные на группы

Funerary inscriptions



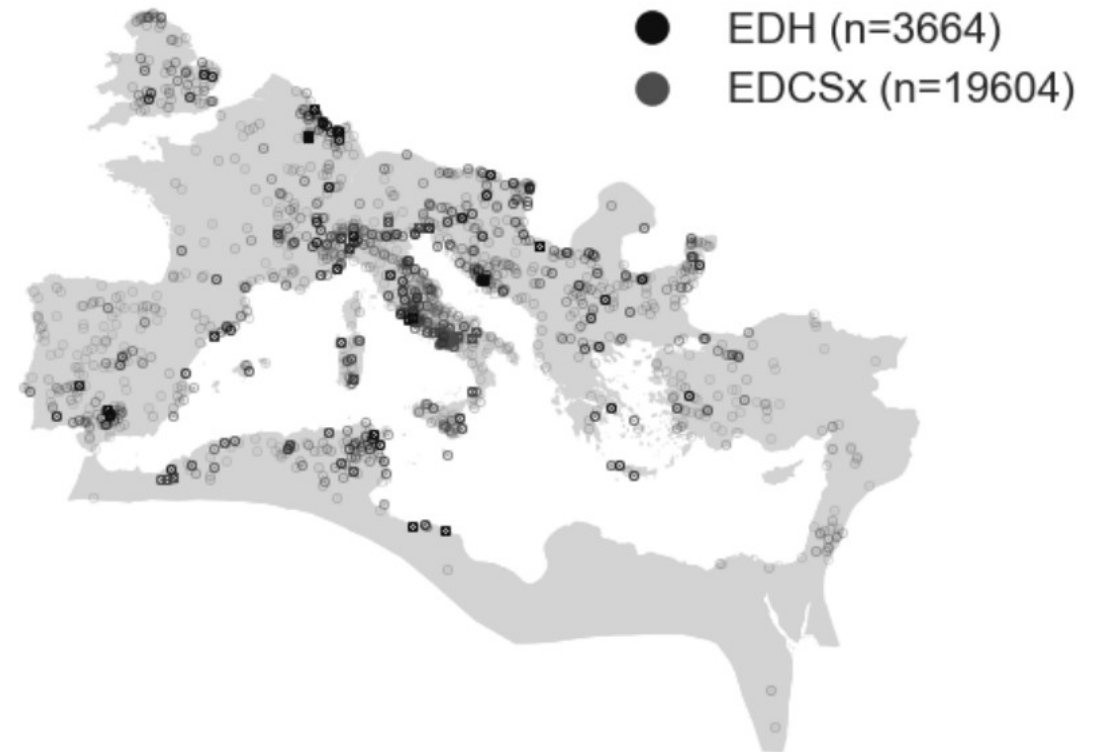
Funerary: specific types



Презентация результатов: карта



Late Roman Empire
(AD 338-AD 600)



Интерпретация результатов: примеры

- Провалы в графике датировок

=> привязка к историческим событиям

- Корреляция близка к нулю

=> это тоже результат

- Распределение на карте

=> видны караванные маршруты

Спасибо за внимание