



## **WOMEN IN DATA SCIENCE**



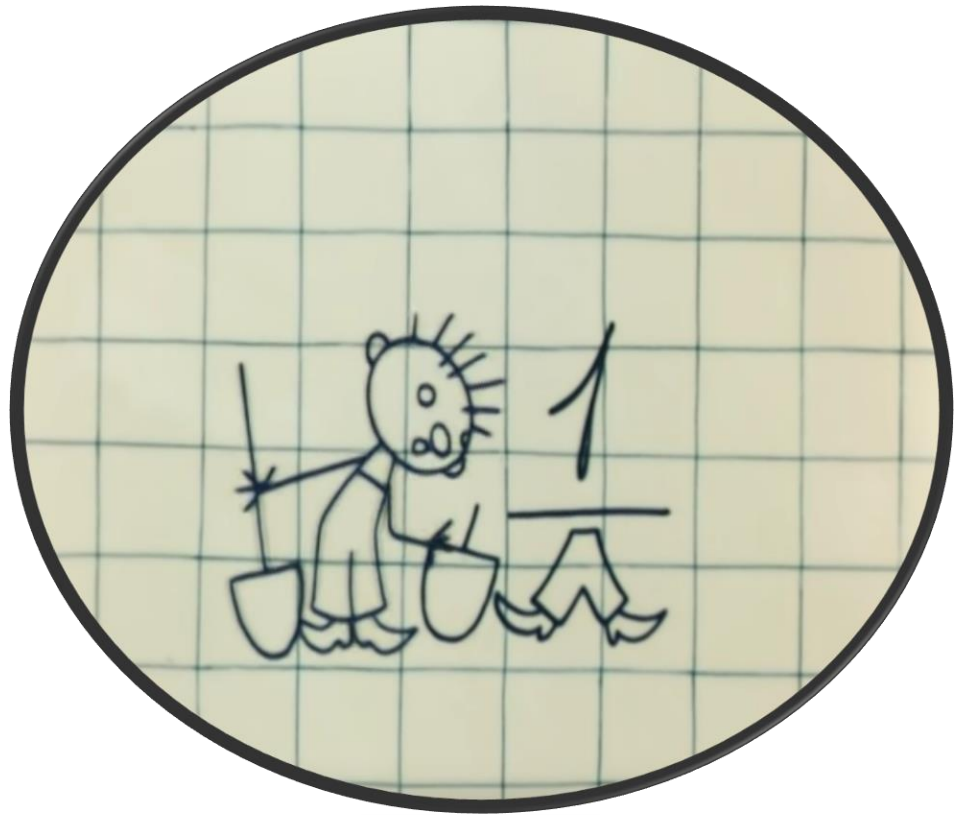
#WiDS2020

# Использование современных технологий в науке о данных (Data science), как альтернатива стандартным методам математической статистики

Федосова Нина  
IBM



# Математическая статистика в современных исследованиях



Группа	Количество животных	Отсутствие осложнений
1	10	70%
2	10	65%
3	10	56.667%
4	10	80%



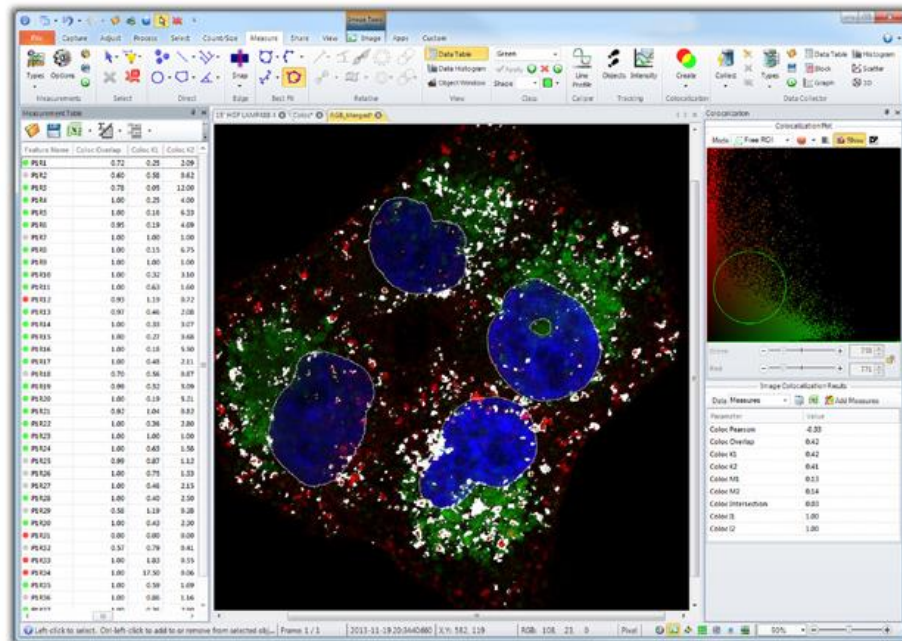
# Необходимые условия применения статистических методов или почему статистика всегда врет

- Нулевая гипотеза об отсутствии статистических различий между группами
- Критический уровень значимости

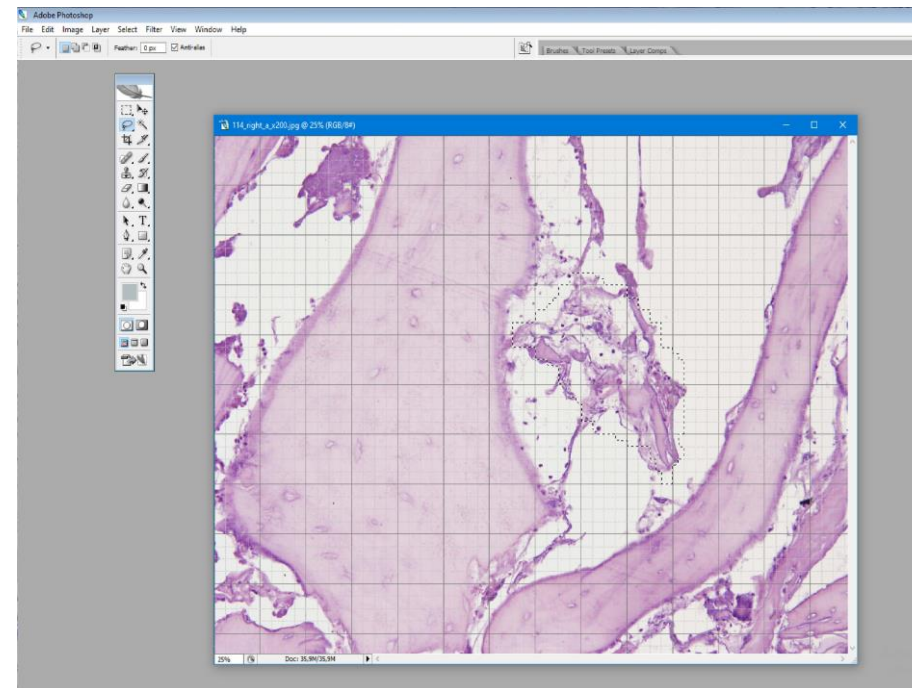


# Современные методы анализа изображений

## Морфометрия

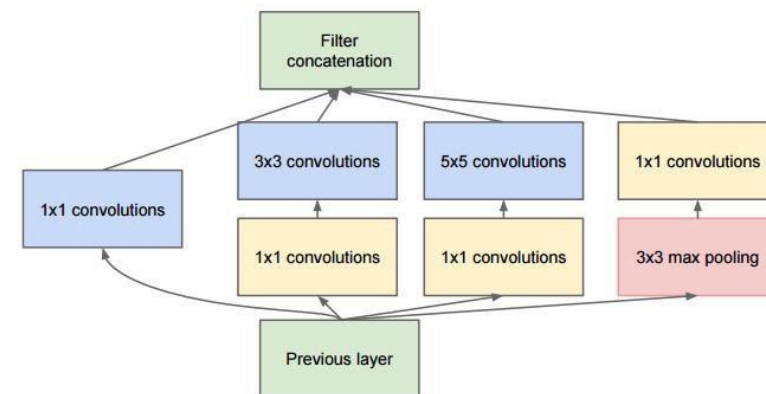
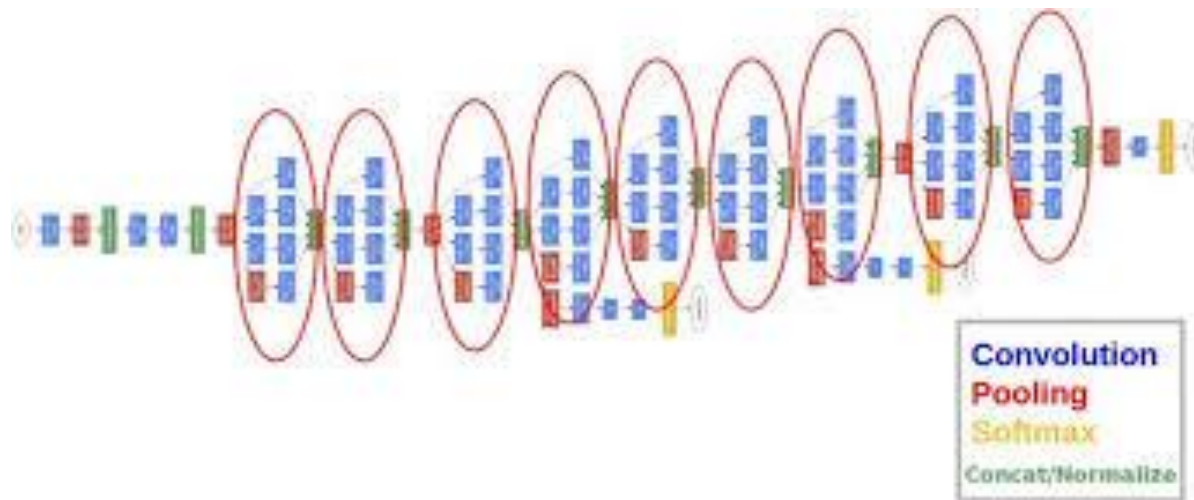


## Photoshop



# Нейронные сети

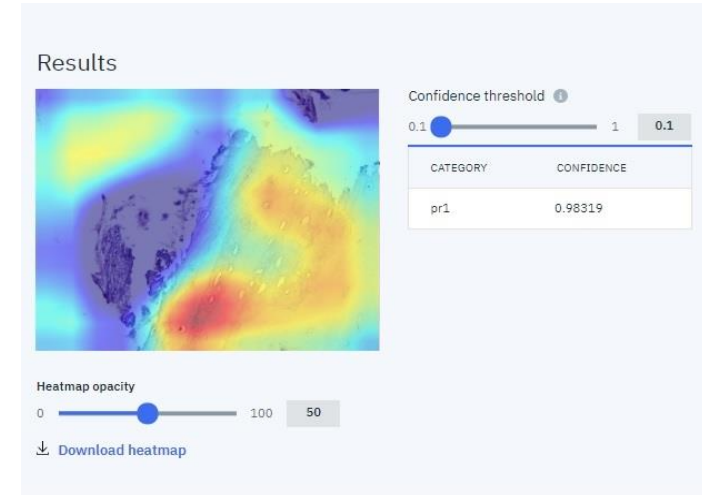
## GoogleNet





# Использование нейронных сетей, как альтернативы методам математической статистики

- Чем определяется статистические различия между сравниваемыми группами
- Оценка результатов эксперимента с помощью созданной модели



# Особенности тонкой настройки нейронных сетей

- Размер набора данных
- Качество изображений
- Размеры изображений
- Использование предварительной обработки изображений
- Создание алгоритма обучения нейронной сети





# Архитектура нейронной сети

- Проблема исчезающего или взрывающегося градиента
- Архитектура современных нейронных сетей
- Увеличения точности и упрощение сетей



# Заключение и выводы

- Выбор архитектуры в соответствии с требованиями к обработке данных
- Использование от 5 до 10 тысяч изображений для обучения
- Отсутствие скачков или затухания метрик при обучении.
- Результаты функционально тестирования на наборе данных более 100 изображений дают единичное число ошибок (меньше 5)



# Вопросы

?

