

# ОСНОВНА МАТЕМАТИКА

ДАЈТЕ СИ ШАНСА!



The “Accountability, Technology and Institutional Openness Network in South East Europe - ACTION SEE” project is implemented by Metamorphosis Foundation, Westminster Foundation for Democracy, CRTA – Center for Research, Transparency and Accountability, Citizens Association Why not?, Center for Democratic Transition, Open Data Kosovo (ODK) and [Levizija Mjajt!](#).

This project is funded by the European Union

# Во оваа презентација

Сериозно, дајте си шанса

- Малку статистика
- Стапки, просеци, средни вредности итн

# ШТО?!

- Калкулации со користење: проценти, медијана, стапки, просеци
- Користење на стандардната девијација за идентификување на отстапувања или outliers
- Корелации!

# Очекувајте шеми/патерни

Тие се во секој сет на податоци

- Анализата на податоците е всушност пронаоѓање шаблони
- Наоѓање докази за нешто невообичаено
- Пронаоѓање на невообичаеното, тоа што „отскокнува“

# Добра вест

Не е тешко

- Лошата вест е дека евентуалните грешки го уништуваат кредибилитетот
- Станува збор за само четири операции: собирање, вадење, множење и делење

# Подготвени?

Да започнеме!



# Процентуална промена

## Формулата на нинџите

- Споредба на **нов** број со **стар** број
- **Формула: (НОВ - СТАР) / СТАР**

**Пример: 24 убиства во 2014 год .. па 30 во 2015 година**

Која е процентуалната промена?

# Стапки

## Споредба на групи со различни големини

- На пример, кога имате градови со различен број на популација
- **Формула: НАСТАНИ / ПОПУЛАЦИЈА \* по единица**

Пример:

**Град 1: 75.000 популација - 40 смртни случаи**

**Град 2: 115.000 популација - 55 смртни случаи**

Во кој град стапката на смртни случаи на 10.000 жители е повисока?



# Униваријантна анализа

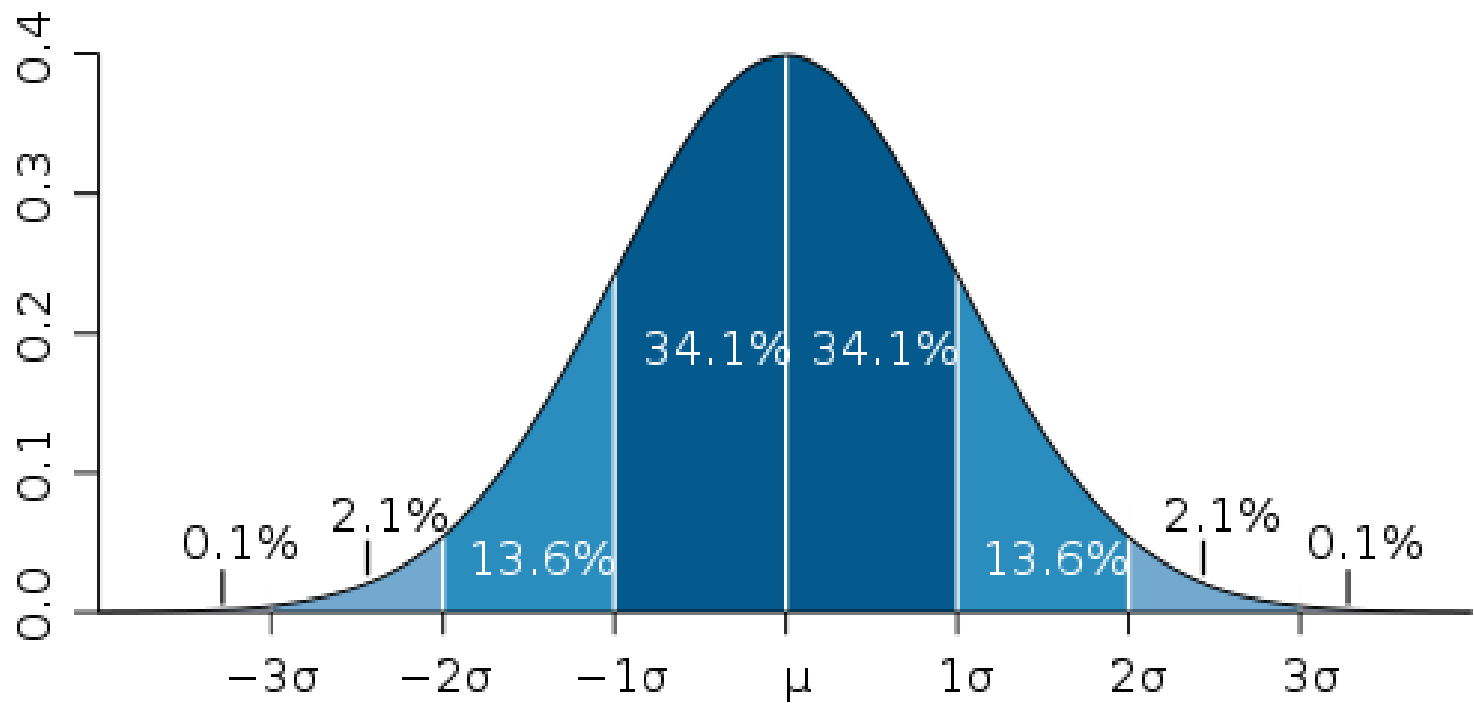
Во шоуто има само една ѕвезда

- Една варијабла во еден сет на податоци
- Дескриптивна статистика: ги опишуваме нејзините карактеристики користејќи калкулации

# Мерки кон центарот

## Наоѓање на средината

- **Просек:** сума на вредностите поделено со бројот на вредности
- **Медијана:** вредноста во средината на една подредена листа
- **Модалитет:** најчестата вредност во една група
- **Отстапувања/Outliers:** вредности кои се далеку од просекот
  - **ПРЕДУПРЕДУВАЊЕ:** Најчесто овде се наоѓа најважното



Нормална дистрибуција

# Нормална дистрибуција

Позната и како “The Bell Curve” (крива-свонче)

- Врвот е во **средината**, кај **медијаната**
- Што подалеку од просекот, тоа поголема е стандардната девијација
- Кривата содржи 100% од податоците

**Пример: Платите на фудбалерите**

# Варијабилност

Податоци кои можат да се менуваат

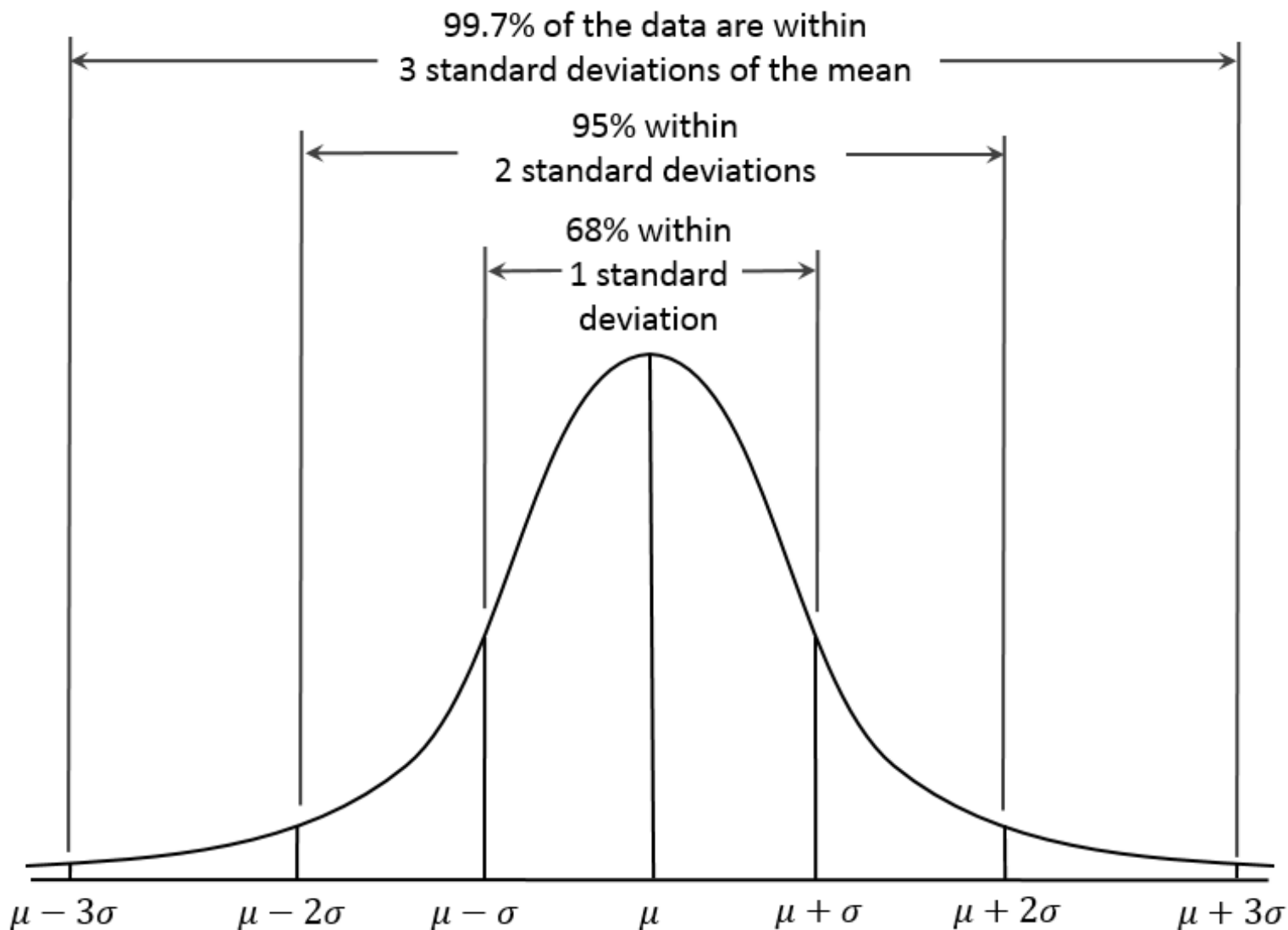
- **Максимум и минимум:** највисоката и најниската вредност во една група
- **Опсег:** разликата меѓу највисоката и најниската вредност во една група
- **Стандардна девијација:** просечната разлика од просекот

# Стандардна девијација

Спасува животи

- Помага да откриеме дали една вредност навистина отстапува/ е **outlier**
- Отстапките **најчесто** се на три стандардни девијации од просекот

# Емпириско правило



# Нормалност

Не се возбудувајте премногу

- Варијабилноста е нормална
- Вредностите во рамки на **ТРИ** стандардни девијации се сметаат за ... нормални.



# Мултиваријантна анализа

## Повеќе луѓе на забавата

- Овде ги правиме корелациите
- **Две или повеќе варијабли**
  - Кога едната оди нагоре, другата оди надолу
  - Кога едната оди нагоре, другата оди нагоре
  - Кога едната оди надолу, другата оди надолу

# Играјте со податоците

- Начин на гледање на работите – analytics mindset
- MIN-MAX-MED view
- **Две или повеќе варијабли?**
  - Логика – common sense: колку е таа навистина “common”
  - Сериозно, само две варијабли
  - Фактори – еколошки, историски, циклични итн.

Пример: поврзете ги бројот на обувки со шансите да се стане пратеник во Собранието на Република Македонија

# Играјте со податоците

FILE HOME INSERT PAGE LA FORMUL DATA REVIEW VIEW DEVELO XL Cam DESIGN Jon Acam...

A3 : X ✓ fx 1/5/2013

	A	B	C	D	E	F	G	H	I	J
1	<b>Source Data for Pivot Table</b>									
2	Date	Qtr	Year	Customer	Region	Product	Quantity	Revenue		
3	01/05/13	Q1	2013	Customer 4	West	Product 9	15	270		
4	03/12/13	Q1	2013	Customer 1	Midwest	Product 3	20	200		
5	03/14/13	Q1	2013	Customer 6	West	Product 8	25	1,150		
6	03/27/13	Q1	2013	Customer 3	West	Product 1	14	100		
7	04/14/13	Q2	2013	Customer 6	Northeast	Product 7	16	400		
8	04/16/13	Q2	2013	Customer 7	Midwest	Product 5	40	510		
9	04/25/13	Q2	2013	Customer 6	South	Product 3	20	70		
10	04/28/13	Q2	2013	Customer 6	Midwest	Product 6	10	92		
11	07/03/13	Q3	2013	Customer 2	West	Product 7	29	350		
12	07/06/13	Q3	2013	Customer 6	Midwest	Product 7	10	128		
13	07/06/13	Q3	2013	Customer 1	Midwest	Product 7	30	660		
14	07/08/13	Q3	2013	Customer 3	West	Product 7	30	276		
15	07/12/13	Q3	2013	Customer 1	Northeast	Product 9	10	530		
16	08/11/13	Q3	2013	Customer 6	West	Product 1	50	500		
17	08/20/13	Q3	2013	Customer 1	South	Product 2	90	2,250		
18	09/14/13	Q3	2013	Customer 3	South	Product 6	50	150		
19	09/23/13	Q3	2013	Customer 4	West	Product 1	25	1,000		
						Product 2	25	75		
						Product 6	15	53		
						Product 3	10	97		

1) Pivot Tables take data...



# БЛАГОДАРАМ



The “Accountability, Technology and Institutional Openness Network in South East Europe - ACTION SEE” project is implemented by Metamorphosis Foundation, Westminster Foundation for Democracy, CRJA – Center for Research, Transparency and Accountability, Citizens Association Why not?, Center for Democratic Transition, Open Data Kosovo (ODK) and [Levizja Mjaft!](#).

This project is funded by the European Union