

ОСНОВНА

МАТЕМАТИКА

ДАЈТЕ СИ ШАНСА!



D.A.T.A.

Data Analytics for Transparency and Accountability

МЕТАМОРФОЗИС 

Фондација за интернет и општество



This project is funded by the European Union.



ZaštoNe?



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 Mjafi!.

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 патерни
- Наоѓање докази за нешто невообичаено
- Пронаоѓање на невообичаеното, исто така

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

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

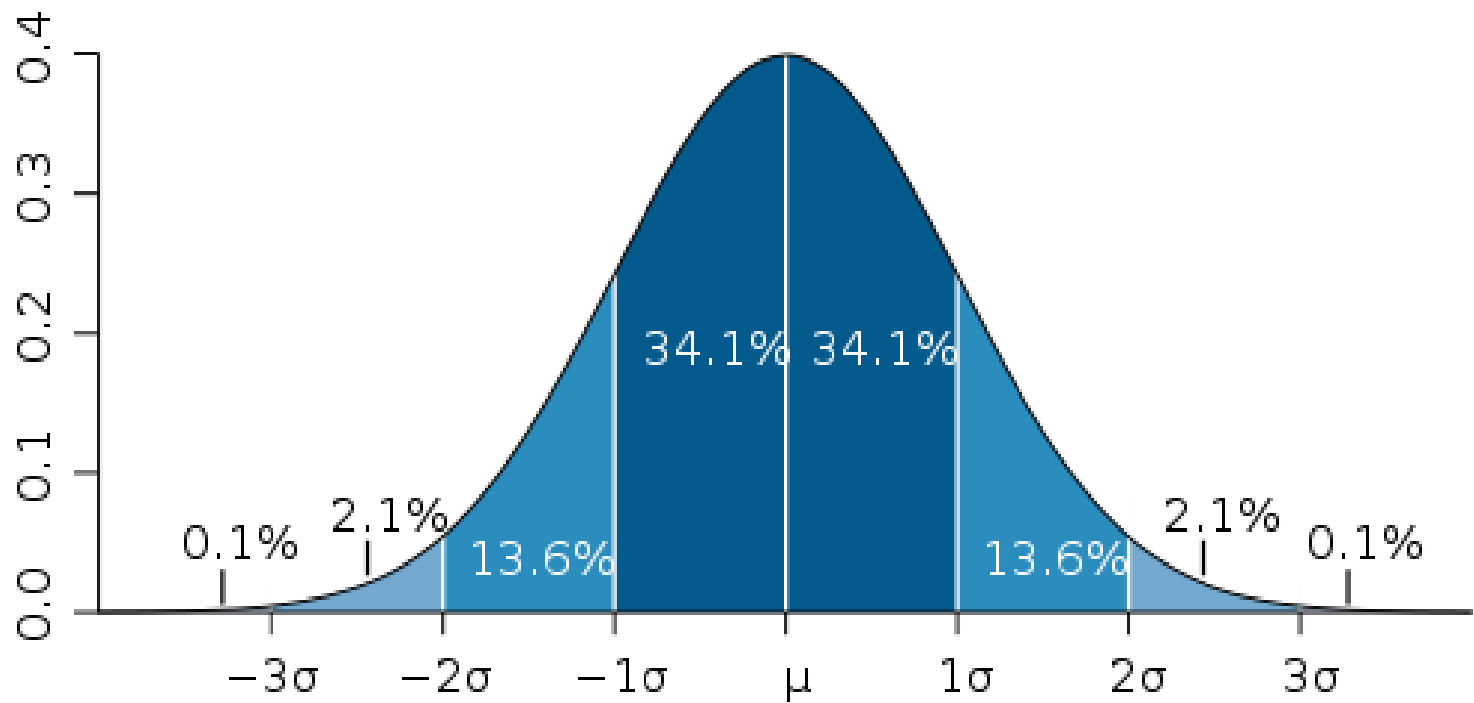
Во шоуто има само една ѕвезда

- 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

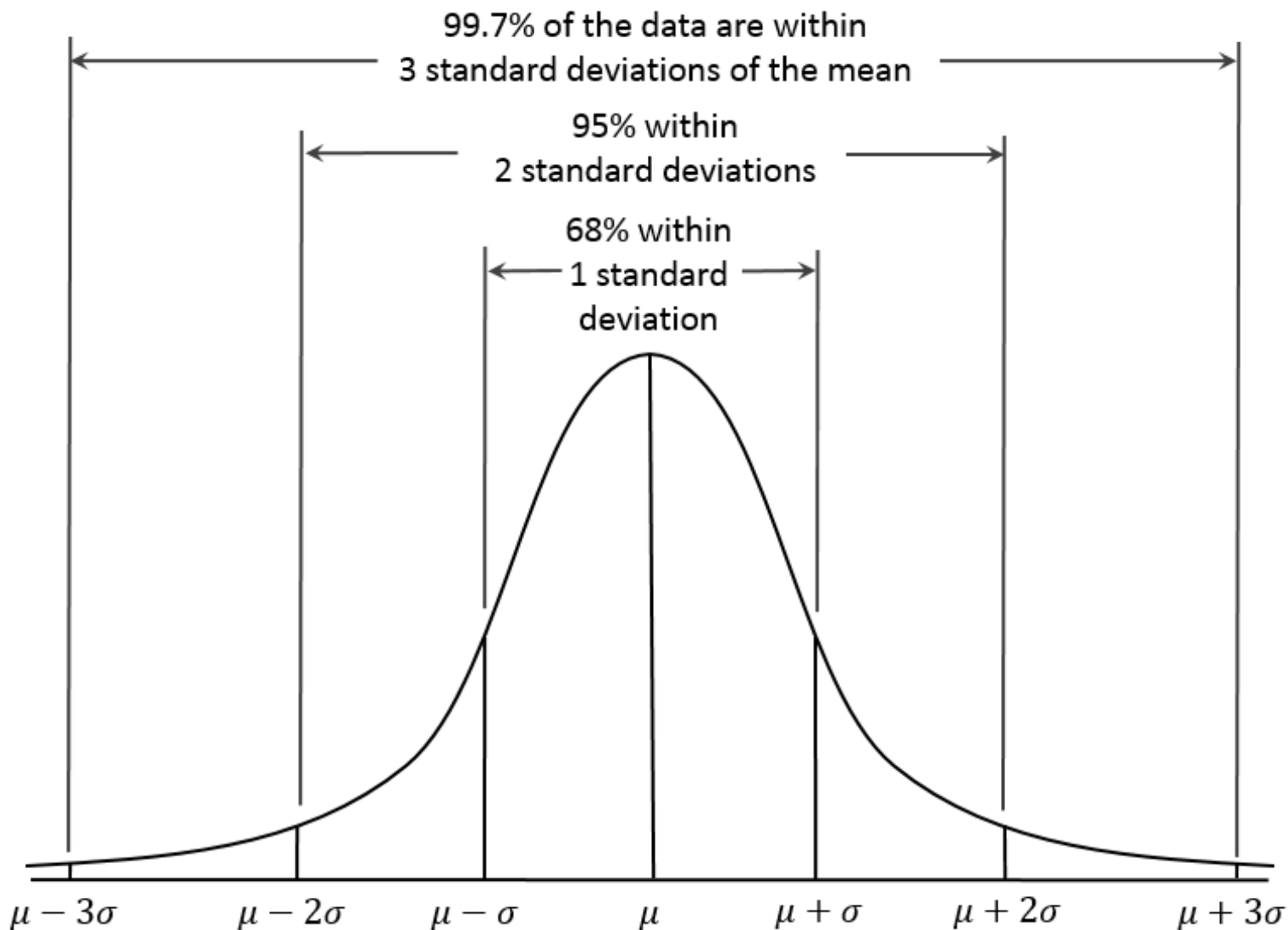
- **Максимум и минимум:** највисоката и најниската вредност во една група
- **Опсег:** разликата меѓу највисоката и најниската вредност во една група
- **Стандардна девијација:** просечната разлика од просекот

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	Source Data for Pivot Table									
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...

Играјте со податоците

Book1 - Micr...

File Home Insert Page Layout Formulas Data Review View

Paste Font Alignment Number Styles Cells

Clipboard Editing

D1 fx

	A	B	C	D	E
1	1	2			Add
2	10	11	12		Range
3	1	2	3		Range
4	5	100			Times

Computer Hope Sheet1

Ready 100%

TotalRow.xlsx - Excel

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW Jeff Lenni...

H61

	A	B	C	D	E	F	G	H	I
6									
7		TransID	AcctNum	Account	DeptNum	Dept	Date	Amount	
8		3630	5056	Meals and	101	Accountin	1/3/2015	854	
12		3634	5054	Internet	101	Accountin	1/30/2015	1,106	
21		3643	5022	Overhead	101	Accountin	4/3/2015	1,308	
31		3653	5051	Small office	101	Accountin	6/12/2015	11	
42		3664	5022	Overhead	101	Accountin	8/28/2015	523	
50		3672	5050	Computer s	101	Accountin	10/23/2015	1,304	
60		Total						5,106	
61									
62									
63									
64									
65									

Total Row

READY 100%

ТОЛКУ!

ДА НУРНЕМЕ



D.A.T.A.

Data Analytics for Transparency and Accountability

МЕТАМОРФОЗИС 

Фондација за интернет и општество



This project is funded by the European Union.



ZaštoNe?



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!