

# Seminář ze statistiky Co je statistika a k čemu ji používáme: aplikace (nejen) v demografii

9. Září 2025

Vysoká škola ekonomická v Praze



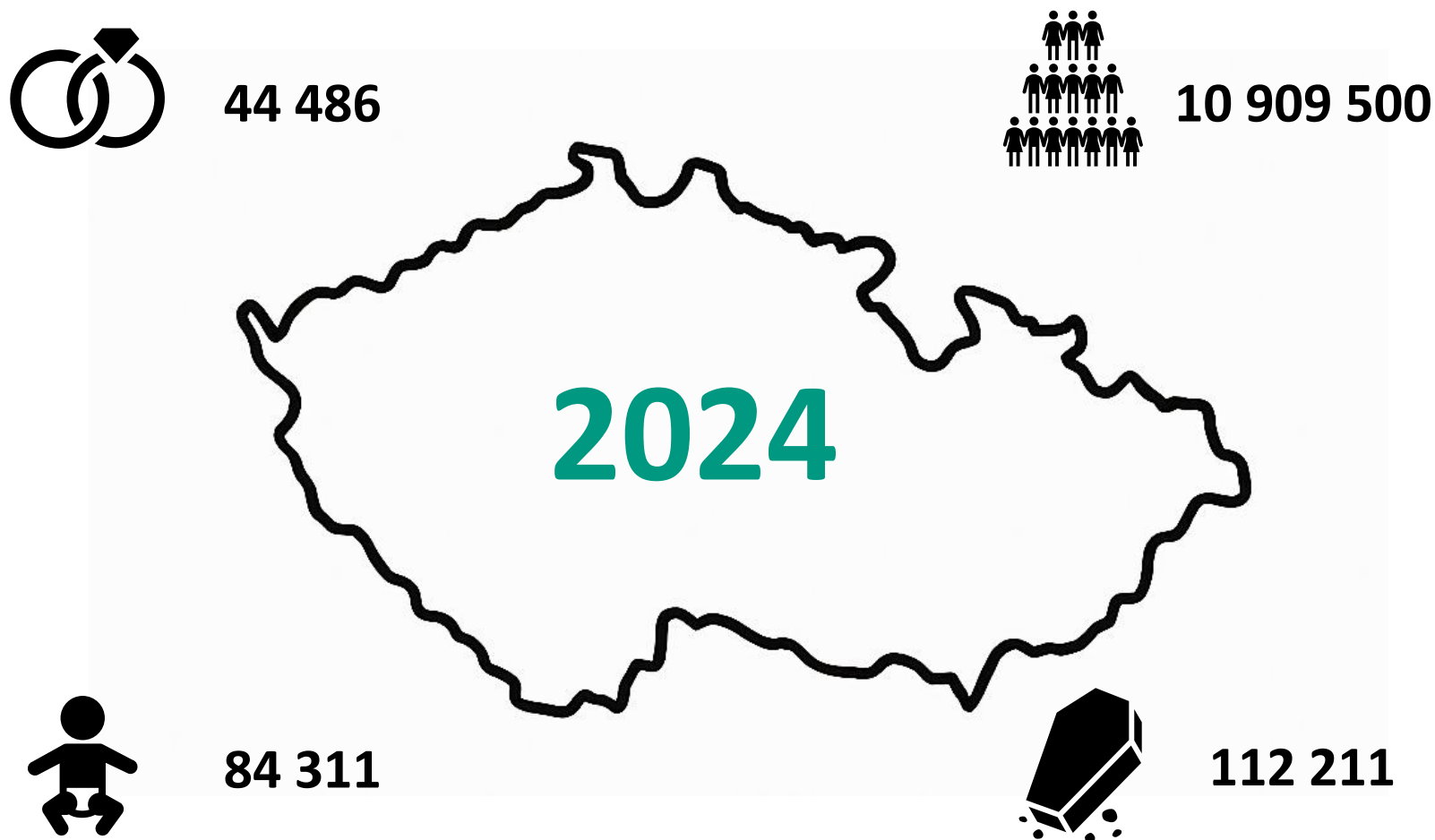
# Obsah semináře – blok 1

1. Co je statistika a proč je důležitá
2. Statistická šetření v praxi (sčítání lidu, porodnost, inflace, volební průzkumy ...)
3. Základní charakteristiky dat (průměr, medián, modus, rozptyl)
4. Praktické aplikace s malými datovými sadami, úskalí průměru a význam rozdělení dat
5. Diskuze: Kde se se statistikou setkáváme v každodenním životě?

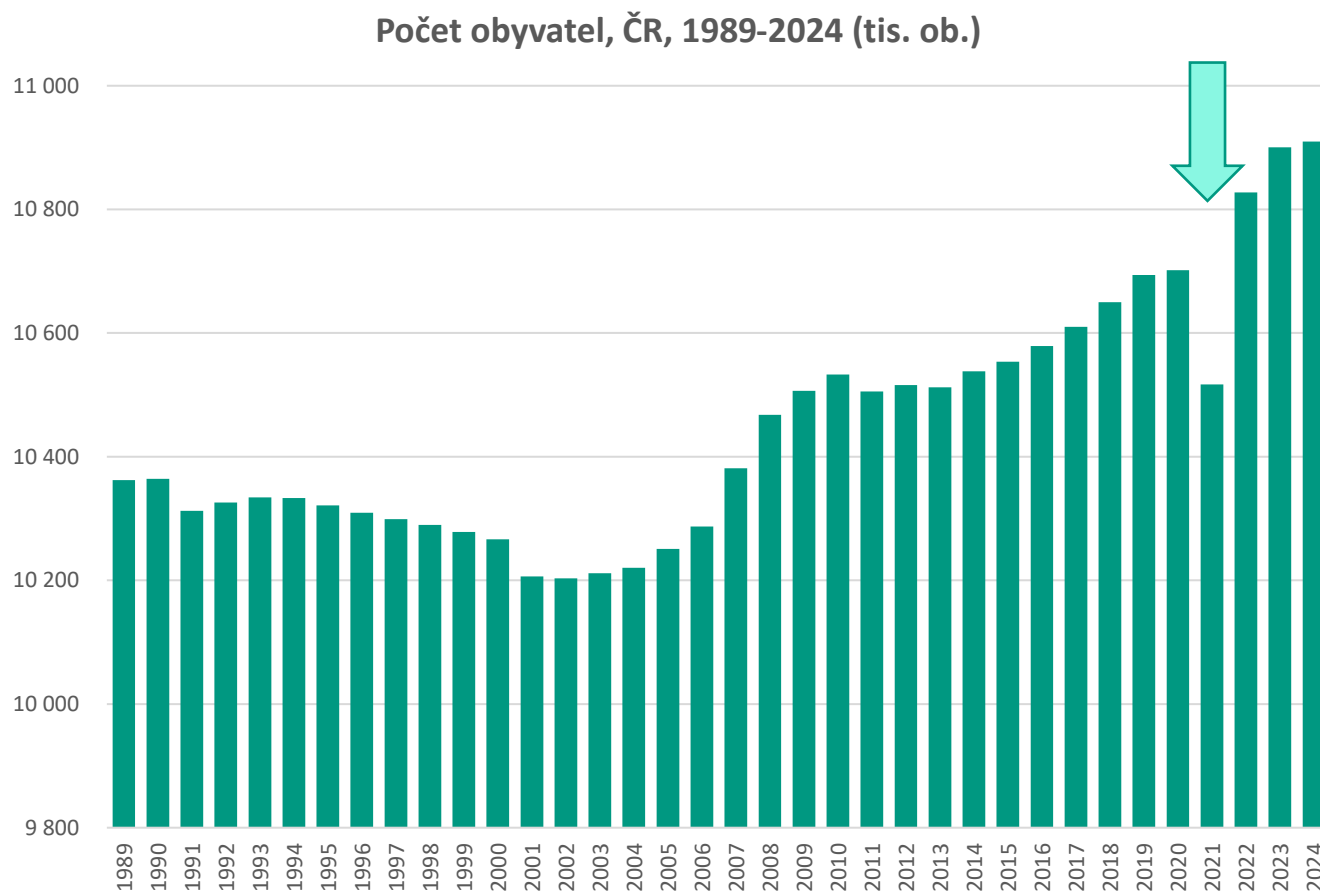
# Statistika a její role v demografii

Vysoká škola ekonomická v Praze

# Statistika a její role v demografii

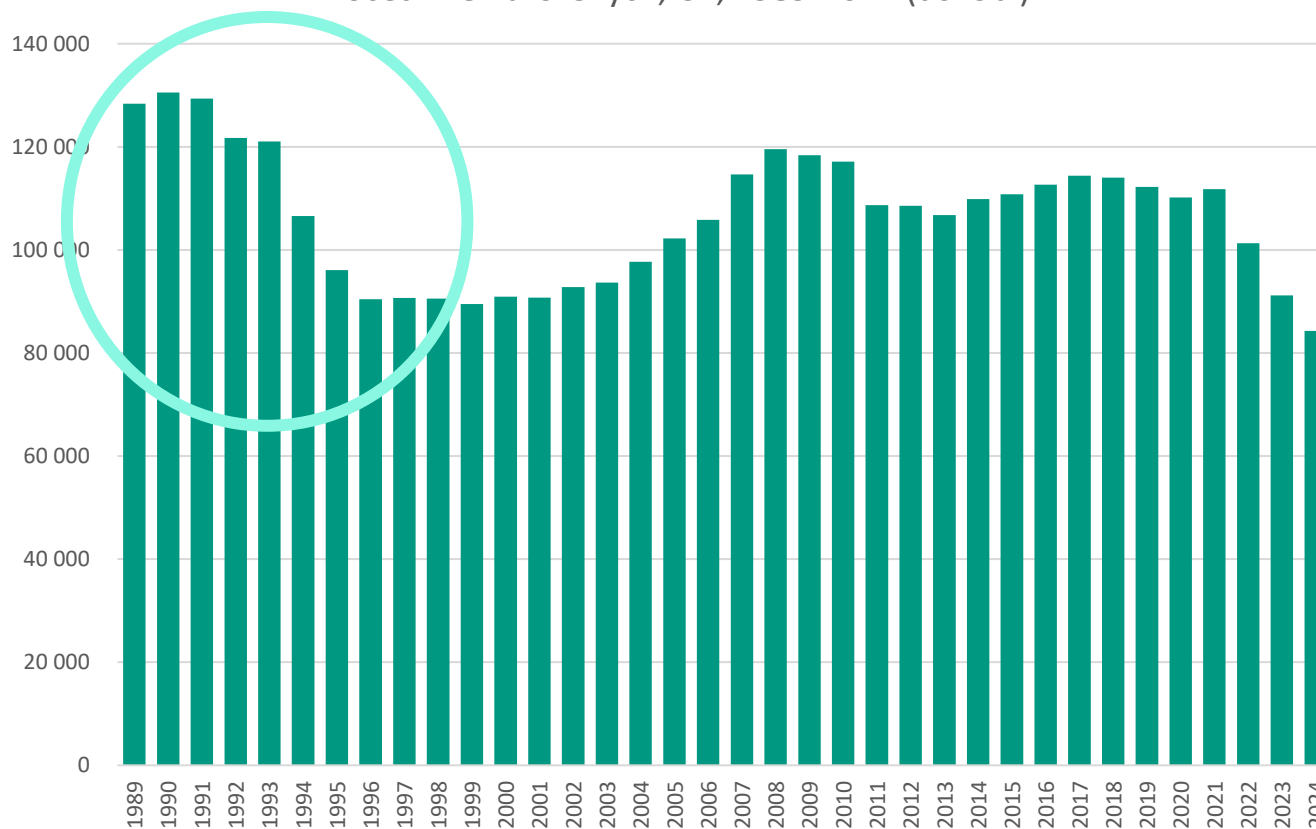


# Statistika a její role v demografii



# Statistika a její role v demografii

Počet živě narozených, ČR, 1989-2024 (tis. ob.)



# Statistika a její role v demografii

**Jak zjistíme, kolik lidí se DNES v ČR narodí?**



# Statistika a její role v demografii

## Jak zjistíme, kolik lidí se DNES v ČR narodí?

### Odpověď nám dávají **statistická šetření**

- Způsob, jak získat informace o větší skupině lidí, věcí nebo jevů

Úplná šetření

Výběrová šetření

#### Český statistický úřad

- Pověřený sběrem, zpracováním a zveřejňováním dat o tom, co se v Česku děje.
- Politicky nezávislý

# Statistická šetření v praxi

Vysoká škola ekonomická v Praze

# Sčítání 2021



- Koná se jednou za 10 let
- Úplné šetření (  výhody  nevýhody)
- Otázky týkající se počtu osob, bydliště, věku, vzdělání, zaměstnání, náboženství ...
- Stát díky tomu může lépe plánovat stavby škol, nemocnic, silnic apod.

# Statistika narození a úmrtí

= tzv. průběžná evidence

- Záznamy z matrik
- Každé narození a úmrtí se hlásí („hlášení o události“)
- Data se zpracovávají průběžné – není to výběr, ale úplný záznam

## Hlášení o narození dítěte

- Datum a místo narození
- Pohlaví
- Živě vs. mrtvě narozené
- Pořadí narození matky
- Věk, rodinný stav, vzdělání matky a otce

## Oznámení o úmrtí

- Datum a místo úmrtí
- Pohlaví
- Věk (datum narození)
- Příčina úmrtí (podle MKN)
- Místo úmrtí (nemocnice, doma, jinde)

# Inflace a spotřební koš

**Inflace** = jak se mění ceny v čase

**Spotřební koš** = soubor „běžných“ věcí a služeb, které lidé nakupují (chléb, mléko, elektřina, veřejná doprava)

- ČSÚ každý měsíc sleduje jejich ceny
- Ukazuje, o kolik zdražuje život a pomocí tohoto ukazatele se následně upravují mzdy a důchody



## Volební průzkumy

- Výběrové šetření: ptají se jen několika tisíc lidí (velké agentury mají svoji reprezentativní základnu respondentů)
- Cílem je odhadnout, jak by dopadly volby, kdyby se konaly teď
- Výsledky se uvádějí s určitou chybou
- Reflektují nálady ve společnosti, pomáhají stranám i voličům

**Jak by volby dopadly dnes?**

# Mini-šetření

Skupinová práce

## Mini-šetření (5 minut)

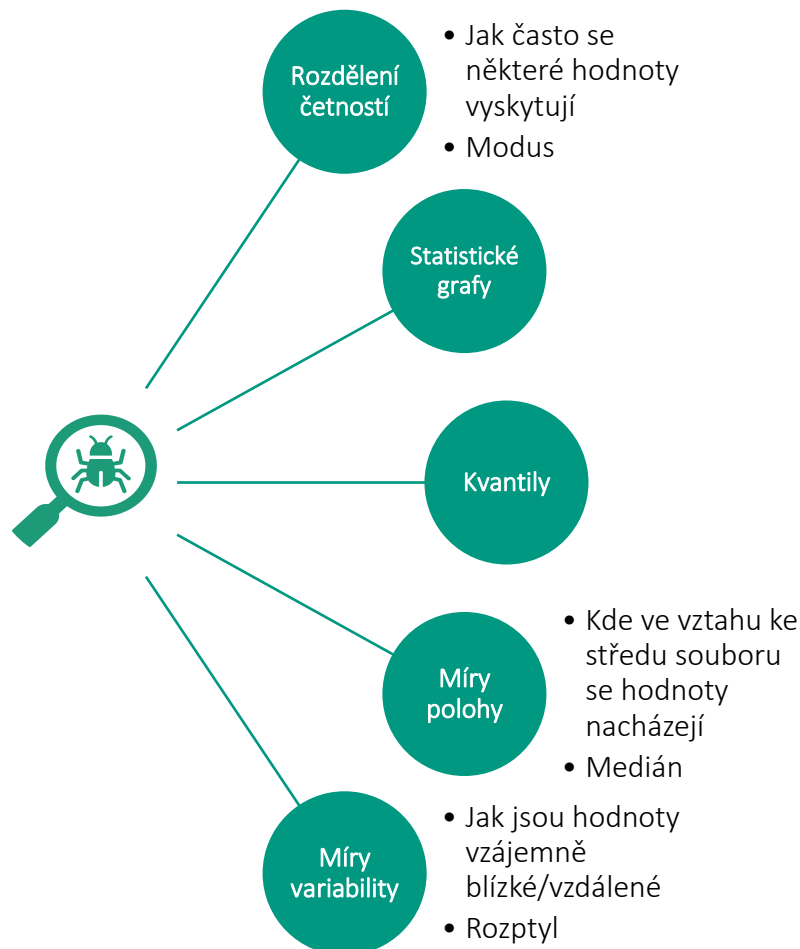
1. Kolik maturantů GJN by chtělo mít v budoucnu 2 a více dětí?
2. Kolik studentů v posledním roce cestovalo do zahraničí?
3. Kolik studentů denně využívá sociální sítě po dobu 3 a více hodin?
4. Jaký je průměrný věk rodičů studentů GJN?



# Základní charakteristiky dat

Vysoká škola ekonomická v Praze

# Základní charakteristiky dat



## Průměr

- „Typická“ hodnota
- Součet všech hodnot dělený jejich počtem

## Medián

- Prostřední hodnota v seřazeném souboru dat

## Modus

- Nejčastější hodnota

## Rozptyl

- Jak moc se jednotlivé hodnoty liší od průměru
- Čím větší rozptyl, tím větší jsou rozdíly mezi hodnotami

# Aritmetický průměr

$\bar{x}$

## PROSTÝ ARITMETICKÝ PRŮMĚR

– pro neseříděná data

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^k x_i}{n}$$

## VÁŽENÝ ARITMETICKÝ PRŮMĚR

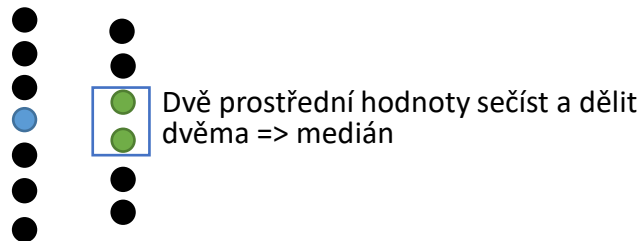
– pro seříděná data (tabulka  
rozdělení četností)

$$\bar{x} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_k n_k}{n_1 + n_2 + \dots + n_k} = \frac{\sum_{i=1}^k x_i n_i}{\sum_{i=1}^k n_i}$$

# Medián

$\tilde{x}$

- hodnoty musíme uspořádat podle velikosti
- každá hodnota se musí vyskytovat tolikrát, kolik je její absolutní četnost



- Funkce **QUARTIL**(pole;kvartil)
  - 1 = dolní kvartil
  - 2 = medián
  - 3 = horní kvartil

Charakterizují  
proměnlivost souboru  
ve stejných měrných  
jednotkách  
jako je statistický znak

## Absolutní míry variability

- **ROZPTYL** = průměr čtverců odchylek jednotlivých hodnot znaku od jejich aritmetického průměru

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \overline{x^2} - \bar{x}^2$$

- **SMĚRODATNÁ ODCHYLKA** = odmocnina z rozptylu

$$s_x = \sqrt{s_x^2}$$

# Praktické aplikace v MS Excel

Vysoká škola ekonomická v Praze

## Příklad 1: Věk matky při narození dítěte

K dispozici jste dostali soubor, který obsahuje věk 144 matek při narození dítěte. Vaším úkolem je vypočítat základní statistické charakteristiky tohoto datového souboru, konkrétně **průměr**, **modus**, **medián** a **rozptyl**. Nejprve proveďte výpočty **ručně podle vzorců**, a poté je **ověřte pomocí funkcí v Excelu**. Zamyslete se nad tím, jak se jednotlivé charakteristiky liší a co nám říkají o rozložení věku matek.

## Příklad 2: Průměrná vs. mediánová mzda

Máme údaje o mzdách pracovníků jedné rodinné firmy (v Kč).

- a) Vypočítejte průměrnou a mediánovou mzdu zaměstnanců.
- b) Jak se změní charakteristiky, pokud paní na úklid (s nejnižší mzdou) dostane přidáno 1 000 Kč?

**Vyskytují-li se v souboru extrémní/odlehlé hodnoty, může medián podat relevantnější informaci o „typické hodnotě v souboru“ než průměr.**



# Vztah mezi průměrem a mediánem

## 1. Pokud je průměr $\approx$ medián

- Rozdělení je **symetrické**.
- Typicky se jedná o **normální rozdělení** (např. výška lidí).
- Průměr dobře reprezentuje „typickou“ hodnotu.

# Vztah mezi průměrem a mediánem

## 2. Pokud je průměr > medián

- Rozdělení je **pravostranně (kladně) šikmé**.
- V datech jsou **vyšší extrémní hodnoty**, které „táhnou“ průměr nahoru.

**Příklad:** Mzdy ve firmě, kde pár manažerů má výrazně vyšší plat než ostatní.

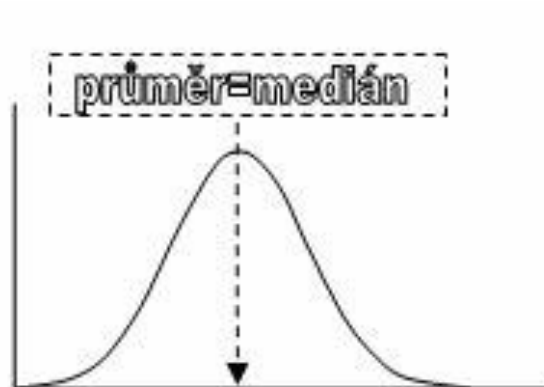
# Vztah mezi průměrem a mediánem

## 3. Pokud je průměr < medián

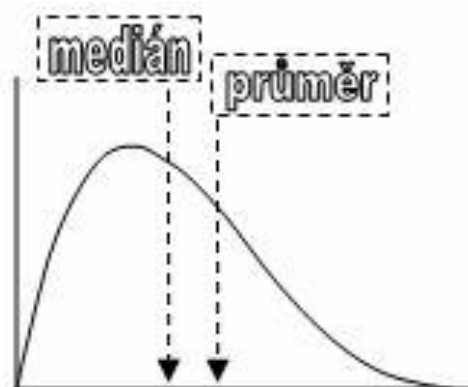
- Rozdělení je **levostranně (záporně) šikmé**.
- V datech jsou **nižší extrémní hodnoty**, které „táhnou“ průměr dolů.

**Příklad:** Výsledky testu, kde většina studentů uspěla, ale pár mělo velmi nízké skóre.

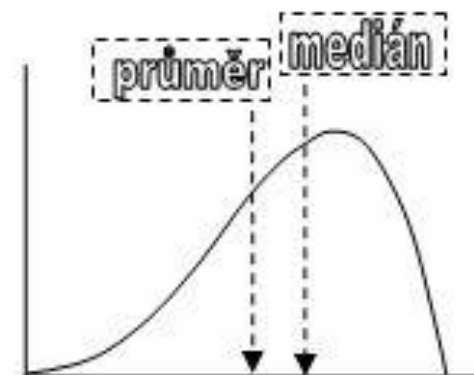
# Vztah mezi průměrem a mediánem



(a) Symetrické  
(nezešikmené) rozdělení



Kladně zešikmené  
rozdělení



Záporně zešikmené  
rozdělení

# Shrnutí a diskuse

Vysoká škola ekonomická v Praze

## Otázky k zamyšlení

- Kde v životě narážíte na „průměrného Čecha“
- Co je problém s průměrnými daty?
- Proč stát, ekonomové nebo demografové potřebují znát nejen průměr, ale i rozdělení dat?

# Děkuji za pozornost

Vysoká škola ekonomická v Praze