

Příklad

Zadání

Ohodnotit učitele podle toho, jak moc učí, vede práce, publikuje...

(Cílem NENÍ program, cílem je získat určité informace; program je jenom prostředek.)

Zadání

PROBLÉM:

Co je málo a co je hodně?

A jak se dá porovnávat třeba vědecká práce a učení?

Zadání

PROBLÉM:

Co je málo a co je hodně?

A jak se dá porovnávat třeba vědecká práce a učení?

ROZHODNUTÍ:

Spočítat pořadí učitele podle různých dostupných kritérií.

Zadání

PROBLÉM:

Jak určovat pořadí v případě SHODNÝCH hodnot?

Zadání

PROBLÉM:

Jak určovat pořadí v případě SHODNÝCH hodnot?

MOŽNOSTI:

- A) interval 10-20
- B) rozhodnout podle dalšího kritéria
- C) střed intervalu
- D) ...

Zadání

PROBLÉM:

Jak určovat pořadí v případě SHODNÝCH hodnot?

ROZHODNUTÍ:

Uvádět pořadí prvního ze všech se stejnou hodnotou.

Zadání

OTÁZKA:

Jaká kritéria?

Zadání

OTÁZKA:

Jaká kritéria?

ROZHODNUTÍ:

počet vyučovaných hodin týdně

počet zápočtů, počet zkoušení

hodnocení ve studentské ANKETĚ

počet vedených a obhájených bakalářských prací

počet oponovaných bakalářských prací

počet vedených a obhájených diplomových prací

počet oponovaných diplomových prací

počet vedených a obhájených disertačních prací

počet publikací

počet publikací, které mají doi, IF, jsou ve WoS

Zadání

PROBLÉM:

ROZVRH, ANKETA a PRÁCE jsou za školní rok.
PUBLIKACE se vykazují za kalendářní rok.

Např. v roce 2026 se vykazují publikace,
které vyšly v roce 2025.

Zadání

PROBLÉM:

ROZVRH, ANKETA a PRÁCE jsou za školní rok.
PUBLIKACE se vykazují za kalendářní rok.

Např. v roce 2026 se vykazují publikace,
které vyšly v roce 2025.

ROZHODNUTÍ:

Výsledek se bude počítat pro školní rok,
PUBLIKACE vyšlé v roce 2025 se budou počítat
do školního roku 2024/25

(důvod: většina roku 2024/25 byla v roce 2025).

Dekompozice

POSTŘEH:

Získání dat bude pomalé (stahovat z webu)
a přitom jednou získaná data se nemění

ROZHODNUTÍ:

Oddělit získávání dat od jejich zpracování
(a to se možná bude dělit dál).

Jak získat data

SIS počet vyučovaných hodin týdně

X počet zápočtů

X počet zkoušení

SIS hodnocení ve studentské ANKETĚ

SIS počet vedených a obhájených bakalářských prací

SIS počet oponovaných bakalářských prací

SIS počet vedených a obhájených diplomových prací

SIS počet oponovaných diplomových prací

SIS počet vedených a obhájených disertačních prací

OBD počet publikací

OBD počet publikací, které mají doi

OBD počet publikací, které mají IF

OBD počet publikací, které jsou indexované ve WoS

Jak získat data

PROBLÉM:

Data o zkouškách a zápočtech nedokážeme jednoduše získat.

ROZHODNUTÍ:

Upřesnění zadání:

Data o zkouškách a zápočtech **vynechat**.

Jak získat data - ROZVRH

počet vyučovaných hodin týdně

- SIS/ROZVRH

- nejde zobrazit pro všechny katedry najednou,
jen po katedrách

= PYTHON, SELENIUM,

stáhne a uloží z webu rozvrh pro každou katedru
a výsledek uloží do textového souboru

= rlistky_2024.txt

= cca 5000 řádek

Jak získat data - ANKETA

SIS/ANKETA

= PYTHON, SELENIUM,

stáhne a uloží z webu výsledky ANKETY pro daný rok

= anketa_2024.txt

= cca 1000 řádek

Jak získat data - PRÁCE

PROBLÉM:

Bakalářská práce, Diplomová práce nebo Disertační práce,
není práce na jeden rok

ROZHODNUTÍ:

Počítat jenom obhájené,
ale za poslední 3 resp. 3 resp. 10 let.

Jak získat data - PRÁCE

- = ručně (protože jednorázový export)
- = SIS/Studentské a závěrečné práce
- = obhájené
- = odkdy-dokdy
- = exportovat do Excelu
- = otevřít a exportovat do CSV,
oddělovač (aby nebyla čárka)
 - = prace_2024.csv
 - = cca 3000 řádek

Jak získat data - PUBLIKACE

PROBLÉM:

Napsat článek trvá dlouho, jeho přijetí a vydání trvá dlouho.

ROZHODNUTÍ:

Počítat publikace za poslední 3 roky.

Jak získat data - PUBLIKACE

= OBD

= filtr - jenom PŘIJATÉ, rok přijetí od-do

= export do Excelu

= otevřít a exportovat do CSV,
oddělovač (aby nebyla čárka)

= Seznamliteratury_2024.csv

= cca 2000 řádek

***** Máme data! *****

Zpracování dat

ÚVAHA:

Potřebujeme zpracovat

ROZVRH, ANKETU, PRÁCE a PUBLIKACE.

A teprve potom, co všechny zpracujeme,
můžeme určit pořadí a budeme moci vypisovat výstup.

Zpracování dat

ÚVAHA:

Potřebujeme zpracovat

ROZVRH, ANKETU, PRÁCE a PUBLIKACE.

A teprve potom, co všechny zpracujeme,
můžeme určit pořadí a budeme moci vypisovat výstup.

ROZHODNUTÍ:

Rozdělit na ZPRACOVÁNÍ a VÝSTUP,

takže budeme mít celkem tři fáze programu:

- 1) získání dat
- 2) zpracování dat = načtení a uložení
- 3) výstup

REPREZENTACE DAT

PODLE ČEHO ROZHODOVAT:

Budeme procházet různá data

a potřebujeme je (rychle) započítávat příslušným učitelům.

REPREZENTACE DAT

PODLE ČEHO ROZHODOVAT:

Budeme procházet různá data
a potřebujeme je (rychle) započítávat příslušným učitelům.

ROZHODNUTÍ:

- Data budou uložena v **DICTIONARY** indexovaném učitelem.
- Budeme do něj ukládat pro jednotlivá kritéria
 - **POČET, SOUČET** (třeba hodiny přednášek a cvičení),
 - **ale i ZDROJOVÁ DATA**, abychom je mohli na konci vypisovat
(když oponoval tři bakalářky, tak vypsát **KTERÉ**)
...takže hodnota v tom dictionary bude mít více položek.

Problém se jmény

Jména

ROZVRH

24aNCGD008x01	X	Praktikum z herního vývoje s řízeným kódem	Holan Tomáš, RNDr., Ph.D.	Černý Vojtěch, Mgr.	Čt 17:20	S8	90
---------------	---	--	---------------------------	---------------------	----------	----	----

RNDr. Tomáš Holan, Ph.D.

ANKETA

Optimalizace [NPRG030, cvičení]	11/65/336	-	1,09:0,30
Programování 1 [NPRG030, cvičení]	9/65/344	-	1,00:0,00
Programování 1 [NPRG030, přednáška]	16/156/344	-	1,44:1,03
Programování 2 [NPRG031, přednáška]	15/137/284	-	1,40:0,83

Analyza a vývoj umělé inteligence založené na Monte Carlo Tree Search pro karetní deskovou hru "7 Divů světa"	2024/2025	bakalářská práce	Matematicko-fyzikální fakulta	Katedra softwaru a výuky informatik (201. • 32-KSVI)	Holan Tomáš, RNDr., Ph.D.
---	-----------	------------------	-------------------------------	--	---------------------------

PRÁCE

TIT Focus: Monitoring and Segmentation of Concentration
AUT Gergelitsová Šárka; Holan Tomáš
ZDR INTED2023 Proceedings, 100-1710
RPT ID 640976 | rok 2023 | může do RIV? Ano | součásti: MFF | spravuje MFF: KSVI | vlastník: Tomáš Holan (MFF) | financování: >I< COOP | RIV/00216208:11320/23:10473315 | C03. Staté a příspěvky cizojazyčné v recenzovaných sbornících
DRU PŘÍSPĚVEK V KONFERENCEČNÍM SBORNÍKU (příspěvek v recenzovaném konferenčním sborníku)

PUBLIKACE

Problém se jmény

PROBLÉM:

Jména jsou zapsaná na různých místech různě.

ROZHODNUTÍ:

- = potřebujeme funkci, která rozhodne, jestli dva různé stringy jsou **tentýž učitel**
- = potřebujeme to dělat **RYCHLE** (bude to mockrát), takže nechceme zkoumat permutace,
- => nějaká kanonická reprezentace (setříděné), a pomocný dictionary osob na překlad jmen (memoizace).
- = i tak to neřeší nejednoznačnost jména bez titulů
 - = V určitých případech nebude správně. **SE STIM SMIŘ!**

Zpracování - ROZVRH

```
[ '23aNAST026p1',
  'P',
  'Dějiny astronomie',
  'Hyklová Petra, RNDr., Ph.D.',
  'Pá 9:00',
  'V125AUUK',
  '180 Sudé týdny (liché kalendářní)',
  'čeština',
  '3',
  'N/1/FAAP-F45`P',
  '',
  '\n' ]
```

0	kod
1	P/x
2	nazev
3	ucitele
4	kdy
5	kde
6	delka
7	jazyk
8	pocet studentu
9	kteri studenti

Zpracování - ROZVRH

PROBLÉM:

Výuka v sudých/lichých týdnech. Jak ji započítávat?

Zpracování - ROZVRH

PROBLÉM:

Výuka v sudých/lichých týdnech. Jak ji započítávat?

ROZHODNUTÍ:

Dělit dvěma.

Zpracování - ROZVRH

PROBLÉM:

"Nepravidelná výuka."

Zpracování - ROZVRH

PROBLÉM:

"Nepravidelná výuka."

ROZHODNUTÍ:

Dělit 10.

Zpracování - ROZVRH

PROBLÉM:

Více učitelů u jedné rozvrhové události.

Zpracování - ROZVRH

PROBLÉM:

Více učitelů u jedné rozvrhové události.

ROZHODNUTÍ:

Dělit rovnoměrně

(i když by bylo možné zjišťovat procenta, kterými se na výuce podílejí).

Zpracování - ANKETA

OTÁZKA:

Když je učitel hodnocený z více předmětů
nebo PŘ a CV stejného předmětu,
jak počítat jeho hodnocení?

Zpracování - ANKETA

OTÁZKA:

Když je učitel hodnocený z více předmětů
nebo PŘ a CV stejného předmětu,
jak počítat jeho hodnocení?

ROZHODNUTÍ:

Medián

(a když je SUDÝ počet, tak průměr dvou prostředních).

Zpracování - ANKETA

RNDr. Tomáš Holan, Ph.D.

Algoritmizace [NPRG062, cvičení]

7/43/346 -

1,29:0,49

Programování 1 [NPRG030, cvičení]

5/46/352 -

1,40:0,89

Programování 1 [NPRG030, přednáška]

20/147/352 -

1,30:0,66

...

Zpracování - PRÁCE

ROZHODNUTÍ:

Každá BP a DP se započítává vedoucímu a oponentovi, DISP jenom vedoucímu.

(Důvod: Oponenti DISP bývají často externisté z jiných škol.)

Zpracování - PRÁCE

0	Název	8	Zájemci
1	Rok vypsání	9	Řešitel
2	Typ	10	Oponenti
3	Fakulta	11	Expiruje
4	Ústav	12	Stav
5	Vedoucí	13	Datum zadání
6	Obor práce	14	Datum obhajoby
7	Obor řešitele	15	Rok obhájení

Zpracování - PUBLIKACE

PROBLÉM:

Publikace jsou různě hodnotné.

Zpracování - PUBLIKACE

PROBLÉM:

Publikace jsou různě hodnotné.

ROZHODNUTÍ:

Rozlišovat čtyři typy publikací:

- 1) všechny
- 2) ty, které mají doi
- 3) ty, které mají IF
- 4) ty, které jsou ve WoS

Zpracování - PUBLIKACE

1	ID publikace	16	Ročník	31	Umístění
2	Do RIV?	17	Číslo	32	Dostupnost
3	Neuplatněný	18	ISSN	33	Název anglicky
4	Stav	19	ISBN	34	Anotace anglicky
5	Druh	20	Náz.zdr.zkráceně	35	Kl. slova anglicky
6	Poddruh	21	IF	36	Název konference
7	Jazyk	22	IF	37	Datum konání - začátek
8	Název	23	Impakt faktor rok	38	Datum konání - konec
9	Anotace	24	Vydavatel	39	Místo konání
10	Počet tvůrců	25	Vydání	40	Stát
11	Autoři	26	Strany	41	Číslo článku
12	Rok	27	Lokalizace	42	DOI
13	Název zdroje	28	Počet stran výsledku	43	UT-WoS
14	Typ zdroje	29	Náklad	44	Financování
15	Místo vydání	30	Primární obor	45	Datum vložení

Zpracování - PUBLIKACE

PROBLÉM:

Publikace často mívají více autorů.

Zpracování - PUBLIKACE

PROBLÉM:

Publikace často mívají více autorů.

ROZHODNUTÍ:

Započítávat každému autorovi podíl ($1/\text{počet_autorů}$).

Zpracování - PUBLIKACE

PROBLÉM:

Dělením počtem autorů může vzniknout nekonečná perioda.

Zpracování - PUBLIKACE

PROBLÉM:

Dělením počtem autorů může vzniknout nekonečná perioda.

ROZHODNUTÍ:

Zaokrouhlovat na dvě desetinná místa.

Výstup

PROBLÉM:

Chceme pro každého určit **POŘADÍ**,
ale to je příliš podrobné (a málo-vypovídající).

Výstup

PROBLÉM:

Chceme pro každého určit **POŘADÍ**,
ale to je příliš podrobné (a málo-vypovídající).

ROZHODNUTÍ:

Kromě pořadí určovat i **PERCENTIL** a **KVARTIL**.

Výstup

PROBLÉM:

Z jakého počtu určovat percentil a kvartil?

Když je 2000 učitelů a ročně se obhájí 500 BP...,
tak všichni budou "A".

Výstup

PROBLÉM:

Z jakého počtu určovat percentil a kvartil?

Když je 2000 učitelů a ročně se obhájí 500 BP...,
tak všichni budou "A".

ROZHODNUTÍ:

Percentily a kvartily určovat vždycky jenom z těch,
kdo mají v daném kritériu nějaká data.

Takže kvartily ABCD a "-".

Výstup

PROBLÉM:

Abychom mohli určit **POŘADÍ** pro jednoho učitele,
musíme určit hodnoty **PRO VŠECHNY**.

Výstup

PROBLÉM:

Abychom mohli určit **POŘADÍ** pro jednoho učitele, musíme určit hodnoty **PRO VŠECHNY**.

ROZHODNUTÍ:

Spočítat to pro všechny učitele.

Výstup

ROZHODNUTÍ:

Formát výstupu:

0) HTML, kvůli snadnému prohlížení a formátování.

1) Pro každé kritérium **USPOŘÁDANÝ SEZNAM**

2) Pro každého učitele

PŘEHLED HODNOT a **POŘADÍ** pro všechna kritéria,

a pro každé kritérium popis zdrojových dat

(třeba vyjmenované hodiny nebo články),

a **GRAF DISTRIBUCE** tohoto kritéria.

Výstup

OTÁZKA:

Jak vytvářet grafy?

Výstup

OTÁZKA:

Jak vytvářet grafy?

ROZHODNUTÍ:

Matplotlib.

Výstup

OTÁZKA:

Kolik sloupečků bude mít (sloupcový) graf?

Výstup

OTÁZKA:

Kolik sloupečků bude mít (sloupcový) graf?

ROZHODNUTÍ:

Pro každý percentil - tj. pokaždé 100 resp. 101.

Důvod:

- * třeba 2000 je moc
- * všechny grafy budou vypadat podobně

Výstup

PROBLÉM:

Jak ukládat vytvořené grafy?

2000 učitelů x 10 kriterií = 20.000 souborů grafů...

Výstup

PROBLÉM:

Jak ukládat vytvořené grafy?

2000 učitelů x 10 kriterií = 20.000 souborů grafů...

ROZHODNUTÍ:

SVG přímo vložené do HTML.

Výstup

PROBLÉM:

Jak určovat VZHLED seznamů i STRÁNEK učitelů?

Výstup

PROBLÉM:

Jak určovat VZHLED seznamů i STRÁNEK učitelů?

ROZHODNUTÍ:

Šablony. Styly.

Výstup

PROBLÉM:

Názvy předmětů jsou někdy hodně dlouhé
a kazí vzhled výstupu.

Výstup

PROBLÉM:

Názvy předmětů jsou někdy hodně dlouhé
a kazí vzhled výstupu.

ROZHODNUTÍ:

Ořezávat, zkracovat.

Výstup

PROBLÉM:

Zkrácením názvu předmětu se někdy ztratí
podstatná informace

"Algoritmy a datové struktury 1"

"Algoritmy a datové struktury 2"

Výstup

PROBLÉM:

Zkrácením názvu předmětu se někdy ztratí podstatná informace

"Algoritmy a datové struktury 1"

"Algoritmy a datové struktury 2"

ROZHODNUTÍ:

Zkracovat tak, aby zbyly poslední čtyři znaky:

24bNPRG038x02 Základy vývoje počítačo... her

PRŮBĚH zpracování a generování

1. Pro každé kritérium:

- * vytvoř seznam učitel-hodnota

- * seříd'

- * každému učiteli urči a ulož jeho pořadí
(a percentil a kvartil).

2. Z těch seznamů vygeneruj seříděné HTML-seznamy

3. Vygeneruj HTML-stránky jednotlivých učitelů.

PRŮBĚH zpracování a generování

PROBLÉM:

Generování stránek je pomalé (cca 2h),
přitom během ladění je potřeba spouštět znovu a znovu.

PRŮBĚH zpracování a generování

PROBLÉM:

Generování stránek je pomalé (cca 2h),
přitom během ladění je potřeba spouštět znovu a znovu.

ROZHODNUTÍ:

1. Ladit Výstup jen na podmnožině učitelů.
2. Během ladění nevytvářet grafy
(jejich vytváření zabírá nejvíce času).

Výsledek návrhu programu

0. Upřesnění zadání

ROZHODNUTÍ:

Spočítat pořadí učitele podle různých dostupných kritérií.

ROZHODNUTÍ:

Uvádět pořadí prvního ze všech se stejnou hodnotou.

ROZHODNUTÍ:

Data o zkouškách a zápočtech vynechat.

ROZHODNUTÍ:

počet vyučovaných hodin týdně

...

počet publikací všech, které mají doi, IF, jsou ve WoS

ROZHODNUTÍ:

Práce počítat jenom obhájené,
ale za poslední 3 resp. 3 resp. 10 let.

0. Upřesnění zadání

ROZHODNUTÍ:

Počítat publikace za poslední 3 roky.

ROZHODNUTÍ:

Z ANKETY budeme zobrazovat Medián všech hodnocení.

ROZHODNUTÍ:

U PUBLIKACÍ započítávat každému autorovi podíl
($1/\text{počet_autorů}$).

Zaokrouhlovat na dvě desetinná místa.

ROZHODNUTÍ:

Kromě pořadí určovat i PERCENTIL a KVARTIL.

0. Upřesnění zadání

ROZHODNUTÍ:

Percentily a kvartily určovat vždycky jenom z těch, kdo mají v daném kritériu nějaká data.

Takže kvartily **ABCD** a "-".

ROZHODNUTÍ:

Výsledek se bude počítat pro školní rok, PUBLIKACE vyšlé v roce 2025 se budou počítat do školního roku 2024/25

(důvod: většina roku 2024/25 byla v roce 2025).

0. Upřesnění zadání

ROZHODNUTÍ:

Formát výstupu:

0) HTML, kvůli snadnému prohlížení a formátování.

1) Pro každé kritérium **USPOŘÁDANÝ SEZNAM**

2) Pro každého učitele

PŘEHLED HODNOT a **POŘADÍ** pro všechna kritéria,

a pro každé kritérium popis zdrojových dat

(třeba vyjmenované hodiny nebo články),

a **GRAF DISTRIBUCE** tohoto kritéria.

ROZHODNUTÍ:

Graf bude mít sloupeček pro každý percentil.

1. Postřehy

2. Algoritmus

1. Pro každé kritérium:

- * vytvoř seznam učitel-hodnota

- * seříd'

- * každému učiteli urči a ulož jeho pořadí (a percentil a kvartil).

2. Z těch seznamů vygeneruj HTML-seznamy

3. Vygeneruj HTML-stránky jednotlivých učitelů.

3. Reprezentace dat

ROZHODNUTÍ:

- Data budou uložena v **DICTIONARY** indexovaném učitelem.
- Budeme do něj ukládat pro jednotlivá kritéria
 - **SOUČET** (třeba hodiny přednášek a cvičení),
 - **ale i ZDROJOVÁ DATA**, abychom je mohli na konci vypisovat
(když oponoval tři bakalářky, tak vypsát **KTERÉ**)
...takže hodnota v tom dictionary bude mít více položek.

ROZHODNUTÍ:

=> pomocný dictionary osob na překlad jmen

4. Dekompozice

ROZHODNUTÍ:

Oddělit získávání dat od jejich zpracování (a to se možná bude dělit dál).

ROZHODNUTÍ:

Rozdělit na ZPRACOVÁNÍ a VÝSTUP, takže budeme mít celkem tři fáze programu:

- 1) získání dat
- 2) zpracování dat = načtení a uložení
- 3) výstup

ROZHODNUTÍ:

= potřebujeme funkci, která rozhodne, jestli dva různé stringy jsou tentýž učitel

5. Diskuze

- * neobsahuje počet zkoušek a zápočtů
- * nedokážeme rozeznat lidi se stejným jménem
- * nepočítáme podíl učitelů na výuce
- * roky pro BP, DP, DISP a PUBLIKACE jsou určeny od oka a mohly by se měnit
- * v ROZVRHU by se mohl započítávat počet studentů
- * ...a v ANKETĚ taky.
- * ty roky BP, DP a DISP by mohly být jiné
- * dvě desetinná místa jsou asi zbytečně přesná

TUPLE

Tuple

Dva různé druhy:

- * referenční System.Tuple
- * hodnotový System.ValueTuple

Tuple jako hodnotový typ

<https://learn.microsoft.com/en-us/dotnet/csharp/language-reference/builtin-types/value-tuples>

```
(double, int) t1 = (4.5, 3);  
Console.WriteLine($"Tuple with elements {t1.Item1} and {t1.Item2}.");  
// Output:  
// Tuple with elements 4.5 and 3.
```

```
(double Sum, int Count) t2 = (4.5, 3);  
Console.WriteLine($"Sum of {t2.Count} elements is {t2.Sum}.");  
// Output:  
// Sum of 3 elements is 4.5.
```

Tuple jako hodnotový typ

Jména částí

* Item1, Item2...

* nepovinně uvedená jména - i potom platí Item1, Item2...

Lze dosazovat, i když se části jmenují jinak...

...a po dosazení zůstávají původní jména!

(Dosazují se hodnoty.)

Alias

```
using Rozsah = (int Min, int Max);
```

Tuple jako hodnotový typ

Dekonstrukce

```
int a;  
int b;  
(a, b) = (10, 20);  
(int c, int d) = (30, 40);
```

Porovnávání

```
(int a, byte b) left = (5, 10);  
(long a, int b) right = (5, 10);  
Console.WriteLine(left == right); // output: True  
Console.WriteLine(left != right); // output: False
```

```
var t1 = (A: 5, B: 10);  
var t2 = (B: 5, A: 10);  
Console.WriteLine(t1 == t2); // output: True  
Console.WriteLine(t1 != t2); // output: False
```

Tuple jako referenční typ

<https://learn.microsoft.com/en-us/dotnet/api/system.tuple>

```
var population = new Tuple<string, int, int, int, int, int, int>(
    "New York", 7891957, 7781984,
    7894862, 7071639, 7322564, 8008278);
// Display the first and last elements.
Console.WriteLine("Population of {0} in 2000: {1:N0}",
    population.Item1, population.Item7);
```

nebo

```
var population = Tuple.Create("New York", 7891957, 7781984, 7894862,
    7071639, 7322564, 8008278);
```

Deklarace proměnných pomocí var

<https://learn.microsoft.com/en-us/dotnet/csharp/programming-guide/classes-and-structs/implicitly-typed-local-variables>

Nemusíme popisovat typ,
odvodí se z výrazu na pravé straně dosazení:

```
// i is compiled as an int  
var i = 5;
```

```
// s is compiled as a string  
var s = "Hello";
```

```
// a is compiled as int[]  
var a = new[] { 0, 1, 2 };
```