

```
1 -- Projekt SQL
2
3 -- tvorba tabulky t_Jan_Pospisil_project_SQL_primary_final
4
5
6 -- zkoumání zdrojových dat
7 select *
8 from czechia_price cp ;
9
10 select *
11 from czechia_price_category cpc ;
12
13 select *
14 from czechia_price cp
15 group by category_code; -- 27 položek - odpovídá czechia price category code
16
17 select *
18 from czechia_payroll cp;
19
20
21 select *
22 from czechia_payroll_calculation cpc ; -- 100 fyzický
23                                     -- 200 přepočtený
24
25 select *
26 from czechia_payroll_industry_branch ; -- code A - zemědělství, B - Těžba a dobývání, C.. , D.., atd.... 19 položek
27
28
29 select *
30 from czechia_payroll_unit; -- pouze jednotka 200 Kč , 80 403 tis. osob, 2 položky
31
32
33 select *
34 from czechia_payroll_value_type; -- 316 Průměrný počet zaměstnaných osob
35                                -- 5 958 Průměrná hrubá mzda na zaměstnance
36
37
38 select *
39 from czechia_payroll cp
40 where value is not null
```

```
41      and value_type_code = 5958
42      and unit_code = 200;
43
44 -- dále potřeba udělat průměr hodnot z jednotlivých kvartálů pro každý rok
45
46 create or replace view jan_pospisil_avg_pay_by_year AS
47 SELECT
48     id ,
49     round(AVG (value), 0) as avg_pay,
50     industry_branch_code ,
51     payroll_year
52 from czechia_payroll cp
53 where value is not null
54     and value_type_code = 5958
55     and unit_code = 200
56     and calculation_code = 100
57     and industry_branch_code is not null
58 group by payroll_year, industry_branch_code
59 order by industry_branch_code, payroll_year;
60
61 select *
62 from jan_pospisil_avg_pay_by_year
63 order by industry_branch_code, payroll_year;
64
65 -- testování lagged
66
67 SELECT
68     id ,
69     LAG(avg_pay) OVER (ORDER by industry_branch_code, payroll_year) AS lag_avg_pay,
70     industry_branch_code
71 FROM jan_pospisil_avg_pay_by_year jpapby ;
72
73
74 -- použít partition by industry type ??
75 -- funkce LAG - začátek druhé hodiny - minulá lekce 11
76 -- AVG() (PARTITION BY country ORDER by date rows between 2 preceding and 2 following předchozí 2 následující
77 -- COALESCE (confirmed, 0) AS confirmed -- missing VALUES dej nulu
78
79 -- ÚKOL 1
80 -- tvorba celého scriptu -- výsledkem je tabulka meziročních růstů či poklesů mezd, rok 2000 odfiltrován- nejsou data za 1999, není
```

možné určit meziroční růst.

```
81 -- Z tabulky jsem dále vyřídil jen ty roky a odvětví, které jasně dokazují, ve kterých letech a odvětvích mzdy klesaly.
82
83 -- první select
84 with avg_pay_by_year as (
85     SELECT
86         id ,
87         round(AVG (value), 0) as avg_pay,
88         industry_branch_code ,
89         payroll_year
90     from czechia_payroll cp
91     where value is not null
92         and value_type_code = 5958
93         and unit_code = 200
94         and calculation_code = 100
95         and industry_branch_code is not null
96     group by payroll_year, industry_branch_code
97     order by industry_branch_code, payroll_year
98 ), -- druhý select
99 lagged_avg_pay_by_year AS (
100     SELECT
101         id ,
102         avg_pay,
103         LAG(avg_pay) OVER (ORDER by industry_branch_code, payroll_year) AS lag_avg_pay,
104         industry_branch_code,
105         payroll_year
106     FROM avg_pay_by_year
107 ), -- třetí select - tabulka růstů a poklesů
108 vysledny_rust as (
109     SELECT
110         id,
111         ROUND((avg_pay/lag_avg_pay - 1)*100,1) AS growth_perc,
112         industry_branch_code,
113         payroll_year
114     FROM lagged_avg_pay_by_year
115     where payroll_year != 2000 -- nejsou data za 1999, proto vyřadit
116 ) -- výběr let a odvětví, kdy mzdy klesaly
117 select
118     vr.growth_perc,
119     vr.payroll_year,
```

```
120     cpib.name
121 from vysledny_rust vr
122 JOIN czechia_payroll_industry_branch cpib
123     ON vr. industry_branch_code = cpib.code
124 where growth_perc < 0
125 order by payroll_year, growth_perc;
126
127 -- ODPOVĚĎ - Nedá se říct, že mzdy rostou ve všech odvětvích. V tabulce vidíme, že časté poklesy během sledovaných let jsou v odvětví
    těžby a dobývání.
128 -- v letech 2009 až 2011 a 2020 byly nejčastější poklesy v ubytování a stravování a pohostinství či veřejné správě a obraně. V roce
    2021 se projevuje COVID a
129 -- je znát pokles kulturní a zábavní činnosti.
130
131
132 -- úkol 2
133 -- funkce year
134 SELECT *
135 FROM czechia_price cp
136 WHERE category_code = 114701
137 GROUP BY region_code , date_from ;
138
139 -- partition by - zkouška - nakonec nepoužito
140 -- sesekat na roky
141 SELECT
142     id,
143     AVG(value) OVER (PARTITION BY category_code ORDER BY date_from) AS average,
144     category_code,
145     region_code,
146     year(date_from) as rok_pocatku
147 FROM czechia_price cp
148 WHERE 1=1
149 AND date_from >= '2007-05-01'
150
151 -- jiný způsob... celý script - ODPOVĚĎ NA OTÁZKU výsledná tabulka
152 WITH mleko_chleba AS (
153     SELECT
154         id,
155         round(AVG(value),2) AS average,
156         category_code,
157         year(date_from) AS rok_sberu
```

```
158 FROM czechia_price cp
159 WHERE category_code = 114201 OR category_code = 111301
160 GROUP BY category_code, rok_sberu
161 ),
162 prumerny_plat AS (
163 SELECT
164     id ,
165     round(AVG (value), 0) AS avg_pay,
166     payroll_year
167 FROM czechia_payroll cp
168 WHERE value IS NOT NULL
169     AND value_type_code = 5958
170     AND unit_code = 200
171     AND calculation_code = 100
172     AND industry_branch_code IS NOT NULL
173     AND value_type_code
174     AND payroll_year = 2006 OR payroll_year = 2018
175 GROUP BY payroll_year
176 ORDER BY payroll_year
177 )
178 SELECT
179     pp.*,
180     mc.average,
181     mc.category_code,
182     ROUND(avg_pay/average,0) AS mnozstvi_lze_koupit
183 FROM prumerny_plat pp
184 JOIN mleko_chleba mc
185     ON pp.payroll_year = mc.rok_sberu;
186 -- mléko 114201, chléb 111301
187
188 -- ODPOVĚĎ: V prvním sledovaném období v roce 2006 lze koupit za průměrnou výplatu 1409 litrů mléka a 1262 kg chleba
189 -- V roce 2018 to bylo 1508 litrů mléka a 1233 kg chleba
190
191
192 -- Úkol 3 nezapomenout seřadit !!!!
193 WITH kat_rust AS (
194 SELECT
195     id,
196     round(AVG(value),2) AS average,
197     category_code,
```

```
198         year(date_from) AS rok_sberu
199     FROM czechia_price cp
200     GROUP BY category_code, rok_sberu
201 ),
202 kat_rust_lag AS (
203 SELECT
204     id,
205     average,
206     LAG(average) OVER (ORDER BY category_code, rok_sberu) AS lag_average,
207     category_code,
208     rok_sberu
209 FROM kat_rust
210 ),
211 vysl_neserazeny AS (
212     SELECT
213         average,
214         lag_average,
215         category_code,
216         rok_sberu,
217         ROUND((average/lag_average - 1)*100,1) AS growth_perc
218     FROM kat_rust_lag
219     WHERE rok_sberu != 2006 -- není s čím rovnat;
220 ) -- vysledna tabulka
221 SELECT
222     cpc.name,
223     rok_sberu,
224     growth_perc
225 FROM vysl_neserazeny vn
226 LEFT JOIN czechia_price_category cpc
227     ON vn.category_code = cpc.code
228 GROUP BY category_code, rok_sberu
229 ORDER BY rok_sberu, growth_perc
230 ;
231 -- ODPOVĚĎ: výsledná tabulak seřazena po letech podle kategorií od nejnižšího po nejvyšší růst.
232 -- Vidíme, že např. mezi lety 2006 a 2007 klesla nejvíce cena jablek a nejpomaleji rostl rostlinný tuk
233 -- napoak nejrychleji rostla cena paprik a to o 94,8 %
234
235
236 -- ÚKOL č. 4
237 SELECT *
```

```
238 FROM jan_pospisil_rust_cen;
239
240 SELECT *
241 FROM jan_pospisil_avg_pay_by_year
242 ORDER BY industry_branch_code, payroll_year;
243
244 CREATE OR REPLACE VIEW jan_pospisil_rust_mezd AS -- bez ohledu na odvětví, pouze meziroční růsty
245 WITH avg_pay_by_year AS (
246     SELECT
247         id,
248         round(AVG (value), 0) as avg_pay,
249         payroll_year
250     from czechia_payroll cp
251     where value is not null
252         and value_type_code = 5958
253         and unit_code = 200
254         and calculation_code = 100
255         and industry_branch_code is not null
256     group by payroll_year
257     order by payroll_year
258 ),
259 lagged_avg_pay_by_year AS (
260     SELECT
261         id ,
262         avg_pay,
263         LAG(avg_pay) OVER (ORDER by payroll_year) AS lag_avg_pay,
264         payroll_year
265     FROM avg_pay_by_year
266 )
267 SELECT
268     payroll_year,
269     ROUND((avg_pay/lag_avg_pay - 1)*100,1) AS mzdy_growth_perc
270 FROM lagged_avg_pay_by_year
271     where payroll_year != 2000 -- nejsou data za 1999, proto vyřadit
272
273 -- tabulka meziročního růstu cen potravin bez ohledu na druh potravin, pouze průměrné ceny
274
275 CREATE OR REPLACE VIEW jan_pospisil_rust_cen AS -- bez ohledu na druh zboží
276 WITH kat_rust AS (
277 SELECT
```

```
278         id,
279         round(AVG(value),2) AS average,
280         year(date_from) AS rok_sberu
281     FROM czechia_price cp
282     GROUP BY rok_sberu
283 ),
284 kat_rust_lag AS (
285     SELECT
286         id,
287         average,
288         LAG(average) OVER (ORDER BY rok_sberu) AS lag_average,
289         rok_sberu
290     FROM kat_rust
291 ),
292 vysl_neserazeny AS (
293     SELECT
294         average,
295         lag_average,
296         rok_sberu,
297         ROUND((average/lag_average - 1)*100,1) AS ceny_growth_perc
298     FROM kat_rust_lag
299     WHERE rok_sberu != 2006 -- není s čím rovnat;
300 ) -- vysledna tabulka
301 SELECT
302     rok_sberu,
303     ceny_growth_perc
304 FROM vysl_neserazeny vn
305 GROUP BY rok_sberu
306 ORDER BY rok_sberu;
307
308
309 SELECT *
310 FROM jan_pospisil_rust_cen;
311
312 SELECT *
313 FROM jan_pospisil_rust_mezd;
314
315 -- ODPOVĚĎ NA OTÁZKU výsledná tabulka
316
317 SELECT
```



```
318     jprc.*,
319     jprm.mzdy_growth_perc
320 FROM jan_pospisil_rust_cen jprc
321 LEFT JOIN jan_pospisil_rust_mezd jprm
322     ON rok_sberu = payroll_year
323 WHERE ceny_growth_perc - mzdy_growth_perc > 10
324
325 -- ODPOVĚĎ : Není žádný rok, ve kterém by růst cen překračoval růst mezd o více než 10 procent.
326
327 -- ÚKOL 5
328 create or replace view jan_pospisil_GDP_rust as
329 with czech_GDP as (
330 select
331     country,
332     `year`,
333     GDP
334 from economies e
335 where country = 'Czech Republic' and GDP is not NULL
336 order by `year`
337 ),
338 czech_GDP_lag as (
339 select
340     country,
341     `year`,
342     GDP,
343     LAG(GDP) OVER (ORDER by `year`) AS lag_GDP
344 FROM czech_GDP
345 )
346 SELECT
347     country,
348     `year`,
349     GDP,
350     lag_GDP,
351     ROUND((GDP/lag_GDP - 1)*100,1) AS GDP_growth_perc
352 FROM czech_GDP_lag
353 WHERE `year` != 1990; -- není s čím rovnat;
354
355 -- z předchozího úkolu - růst cen a mezd v letech
356
357 CREATE OR REPLACE VIEW jan_pospisil_mzdy_vs_ceny AS
```

```
358 SELECT
359     jprc.*,
360     jprm.mzdy_growth_perc
361 FROM jan_pospisil_rust_cen jprc
362 LEFT JOIN jan_pospisil_rust_mezd jprm
363     ON rok_sberu = payroll_year;
```

```
364
```

```
365
```

```
366
```

```
367 select*
```

```
368 from jan_pospisil_mzdy_vs_ceny
```

```
369
```

```
370 select*
```

```
371 from jan_pospisil_GDP_rust
```

```
372
```

```
373 -- srovnání
```

```
374 SELECT
```

```
375     jpgr.country,
```

```
376     jpgr.year,
```

```
377     jpgr.GDP_growth_perc,
```

```
378     jpmvc.ceny_growth_perc,
```

```
379     jpmvc.mzdy_growth_perc
```

```
380 FROM jan_pospisil_GDP_rust jpgr
```

```
381 JOIN jan_pospisil_mzdy_vs_ceny jpmvc
```

```
382     on jpgr.year = jpmvc.rok_sberu;
```

```
383
```

```
384 -- ODPOVĚĎ: Výsledné hodnoty by bylo vhodné zobrazit do grafu. Je patrné, že mzdy rostou pomaleji v letech slabších výsledků nebo
poklesů HDP.
```

```
385
```