

Témata pro zkoušku profilové části MZ v předmětu :

DAN - Datová analytika

Školní rok:	2026/2027
Studijní obor:	18 - 20 - M/01 Informační technologie
Zaměření:	Analýza, vývoj a správa
Třída (příp. třídy):	4. F, 4. G, 4. H, 4. I

1. Transformace dat pomocí knihovny Pandas
 - DataFrame, merge
 - formát csv, JSON
 - vizualizace dat
2. Návrh třídy Transform s logováním
 - DataFrame, None, NaN
 - ETL proces, ETL pipeline
 - logging
3. Spuštění dockerizované databáze a extrakce dat
 - docker vs venv prostředí
 - docker compose
 - container, volume
 - princip extrakce dat
4. Otestování transformační funkce pomocí pytest
 - unit testy
 - ověření správnosti výstupu
 - critical path, happy day, alternativní tok
5. Statická analýza a lintování kódu
 - statistická analýza kódu
 - lintování pomocí ruff
 - standard kvality kódu PEP8 a související
6. Profilace výkonu Python skriptu
 - pojem profilování a jeho užití
 - CProfile, timeit
 - různé přístupy k transformaci
7. Čtení dat z API a uložení do souboru
 - REST API, endpoint
 - JWT token
 - knihovna requests

8. Datová analytika ve společnostech
 - typy úloh
 - monetizace
 - životní cyklus projektu DA
 - CRISP-DM
9. Typy dat a složené datové typy
 - kvalitativní x kvantitativní
 - List
 - Tuple
 - Set
 - Dictionary
10. Funkce a moduly v Pythonu
 - definice a použití funkce
 - lambda funkce
 - návratové hodnoty
 - výjimky
11. Numpy
 - ndarray
 - shape, indexace a slicing
 - základní operace s ndarray
 - univerzální funkce
 - náhodná čísla
12. Pandas
 - DataFrame
 - import a export dat
 - práce s řádky a sloupci
 - Indexace, filtrování, změna dat
13. Základní popisné statistiky dat
 - průměry, četnost
 - korelace, kovariance
 - rozptyl, směrodatná odchylka
 - kvartily
 - vizualizace pomocí boxplot
14. Transformace dat
 - DataFrame, spojení více DataFrames
 - normalizace vs standardizace
 - chybějící a odlehlé hodnoty
 - práce s nečíselnými hodnotami
 - regulární výrazy

15. Vizualizace dat

- zásady efektivní vizualizace dat
- popisky os, název, legenda
- druhy grafů
- histogram
- nejčastější chyby při tvorbě grafů

16. Knihovny pro vizualizaci dat

- plotly
- matplotlib
- seaborn
- rozdíly mezi knihovnami
- hierarchie objektů
- interaktivita

17. Úvod do SQL a základní příkazy

- SQL v datové analytice, základní role databázového systému
- SQL DDL, DML, DCL
- struktura příkazu SELECT (projekce, filtrování, řazení)
- DISTINCT, ORDER BY, TOP, OFFSET, stránkování

18. Skupinové dotazy

- význam skupinových dotazů
- GROUP BY
- agregační funkce: COUNT, SUM, AVG, MIN, MAX
- WHERE a HAVING

19. Poddotazy a složené dotazy

- definice poddotazu, WHERE, FROM
- korelovaný a nekorelovaný poddotaz
- efektivita a náročnost poddotazu
- UNION a UNION ALL
- ANY, ALL, EXISTS

20. Práce s výrazy v SQL

- užití CAST, CONVERT a jejich specifika
- CASE
- pivotace pomocí CASE
- řádkové konstruktory

21. Indexy a rychlost dotazů

- význam a užití indexu
- primární, unikátní, složený index
- EXPLAIN, query plan
- vliv indexu na rychlost vykonávání dotazu, důsledky špatně navržených indexů

22. T-SQL a procedury

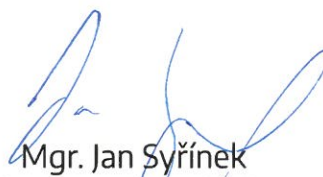
- rozšíření standardního SQL v MS SQL Serveru
- uložená procedura – deklarace, spuštění
- testování a ladění procedur
- provázání událostí s procedurami, trigger, momenty spuštění

23. Zabezpečení v databázích

- význam bezpečnosti v databázích, ochrana dat, zákonné požadavky
- role a uživatelé
- autorizace
- systémová práva a práva uživatelská
- monitorování činností a incidentů (logy, audit)
- odebrání nebo změn oprávnění

Projednáno v katedře VTT a schváleno vedením školy.

V Plzni dne 30. září 2026



Mgr. Jan Syřínek
Zástupce ŘŠ, zástupce statutárního orgánu

Mgr. Jan Syřínek Digitálně podepsal
Mgr. Jan Syřínek
Datum: 2026.09.16
15:12:37 +02'00'