**Příloha č. 10 – Vzorové zadání analytické úlohy**

Modelová analytická úloha pro VŘ na podporu analytiky 2023

Cílem analytické úlohy je poskytnutí přehledu o vývoji bilance čerpacích stanic sítě (dále jen „ČS“) za účelem identifikace možných vad jednotlivých měřidel, případně nelegálních odběrů v rámci logistického řetězce výdej na skladě - přeprava na čerpací stanici (dále jen ČS) - výdej konečnému zákazníkovi.

Tato síť je zásobována jednotlivými zavážkami automobilových cisteren (CAS), na výdejové straně jsou k dispozici záznamy o výdeji. Na jednotlivých stanicích je k dispozici měření objemu v nádržích. Na výdejové straně jednotlivé záznamy odpovídají objemu uvedeném na účtenkách.

**Popis datových zdrojů**

Datovým zdrojem je databáze Microsoft SQL externí firmy. V případě zavážek je objem a druh paliva naložený do CAS deklarován na dodacím nákladovém listě.

**Měření nádrže**

Měření objemu v nádržích je realizováno pomocí měrných tyčí a přidružených měřidel. Perioda měření je zpravidla 20-30 minut (záleží na nastavení na jednotlivých ČS), nicméně dochází k prodloužení periody (a to i v rámci jedné ČS).

Pro každou časovou značku odečtu jsou k dispozici tyto hodnoty:

1. Identifikační číslo stanice
2. Identifikační (pořadové) číslo nádrže
3. Identifikátor paliva
4. Hladina
5. Objem
6. Objem při 15°C
7. Teplota
8. Výška vody

**Výdeje**

Informace o výdejích jsou reprezentovány záznamy přenášenými z účtenek a jednotlivých výdejních stojanů.

Záznam obsahuje následující hodnoty:

1. Identifikační číslo stanice
2. Identifikační (pořadové) číslo nádrže
3. Identifikátor paliva
4. Datum výdeje
5. Datum zápisu záznamu
6. Čas zvednutí pistole
7. Čas zavěšení pistole
8. Objem / Množství (akt. Teplota)
9. Typ účtenky

Při zpracování je nutno brát v úvahu a pracovat s následujícími nestandardními situacemi:

* Neshoda identifikátorů na měření a výdejích
  + Pro dané identifikační číslo nádrže je na výdejích v daném čas pro daný výdej evidováno jiné id paliva tj. v daném čase je a) = i. & b) = ii. Ale c) =/= iii.
  + Nutná dodatečná analýza pro nalezení chybně evidovaného identifikátoru (návrh analýzy + řešení)
* Chybějící časy vi. a vii.
  + Návrh řešení začlenění záznamů s touto vadou do zpracování s minimální chybou

**Zavážky na ČS**

Zavážkami na ČS je míněna dodávka paliva do výdejových nádrží ČS. Dodávka paliva je zaznamenána 1-3 způsoby (dle typu zavážky / lokálního nastavení ČS).

Dodací nákladový list

1. Dodací nákladový list (DNL)

Záznamy dodací nákladových listů reprezentují objem naložený na CAS na daném skladu, každý záznam odpovídá jedné komoře CAS.

Relevantní sloupce záznamu:

* Číslo DNL – identifikátor
* Datum příjezdu CAS na sklad
* Datum odjezdu (resp. vystavení dokladu)
* Jméno řidiče
* Identifikační číslo řidiče
* Objem při 15°C
* Typ paliva (rozdílný číselník od číselníku externí firmy poskytující ostatní údaje)
* Identifikátor destinace – textové pole
* Jméno operátora přítomného při vystavení dokladu

Dataset obsahuje všechny záznamy DNL, je nutno provést analýzu a vybrat pouze ty záznamy které jsou relevantní k úloze (tj. ty které reprezentovaly objem zavezený na ČS) – tj. analýza pole Identifikátor destinace a extrakce identifikačního čísla stanice.

1. Dodávka

Dataset obsahuje záznamy dvojího typu:

1. Dodávky na ČS

Dodávky jsou vytvářeny z nákladových listů (tato data jsou zpracovávána v DB externí firmy v rámci jejich odpovědnosti). Jsou to tedy ty záznamy DNL, které byly na ČS zavezeny, nicméně záznam neobsahuje všechny údaje pro pokrytí všech cílů analytické úlohy.

* Čas vystavení dokladu
* Identifikátor výdejového skladu
* Číslo nákladního listu
* Identifikátor stanice
* Identifikátor paliva (v souladu s Měřením, Výdeji, a bodu III. Zavážek na ČS)
* Objem paliva (při 15°) – jde o sumární objem daného paliva zavezeného na ČS)

1. Převozy PHM z ČS na jinou ČS

V tomto případě obsahuje záznam stejné informace jako Dodávky, s tím že v objemu je uvedeno opačné znaménko. Tyto doklady jsou zadávány manuálně odpovědnou osobou časovou prodlevou (zpravidla 1-3 dny po jejím uskutečněním).

1. Zavážka paliva - měření

Na ČS je vygenerován záznam obsahující údaje z měrných tyčí příjmové nádrže na ČS.

* Záznam obsahuje následující hodnoty:
* Identifikační číslo stanice
* Identifikační (pořadové) číslo nádrže
* Datum a čas začátku plnění nádrže
* Počáteční objem v nádrži
* Počáteční teplota v nádrži
* Datum a čas začátku konce nádrže
* Konečný objem v nádrži
* Konečná teplota v nádrži

**Cíl analytické úlohy:**

Úloha má poskytovat podporu jednak osobám odpovědným za oblast ČS a bezpečnostním analytikům co se týče:

* přehledu denních bilancí jednotlivých ČS a paliv (průběžný inventurní stav dané ČS), denní přehled bilance dané ČS jednotlivých paliv na ní vydávaných a dlouhodobý trend bilance dané ČS v porovnání s celou sítí
* souhrnné bilance ČS za období specifikované uživatelem (podpora pro inventuru ČS)
* detailní přehled vývoje výdejů/příjmů a stavů objemů v jednotlivých nádrží v rámci dne
* bilanci závozů na ČS – tj. porovnání objemů deklarovaných na výdeji s objemy přijatými na ČS
* určení společných jmenovatelů a jejich průběžná prezentace (???) abnormálních odchylek bilancí (v rámci poskytnutého rozsahu poskytnutých dat)

Zpracování úlohy by mělo probíhat plně automaticky v pravidelných intervalech a při zpracování by měla být brána v potaz možná oprava výchozích dat s časovým offsetem až d-20 (tj. daný záznam může být změněn/doplněn/opraven) i s 20ti denním zpožděním.