

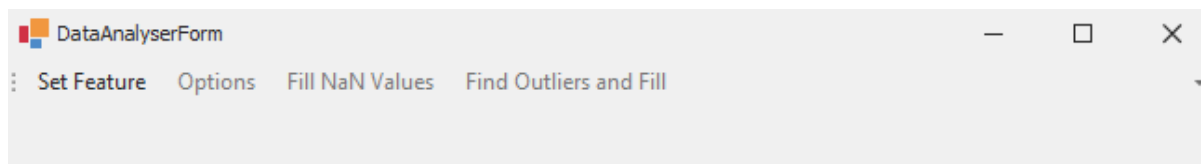
Návod pro ovládání modulu DJ3

Obsah

1. Výběr a filtrace atributu datasetu	2
2. Vizualizace dat	4
2.1. Ovládání grafů	5
3. Metody pro analýzu dat	6
3.1. DJ - KMeans clustering	6
3.2. DJ - SAXpatternDetector	7
4. Značky	9
5. Chybějící a odlehlé hodnoty	10

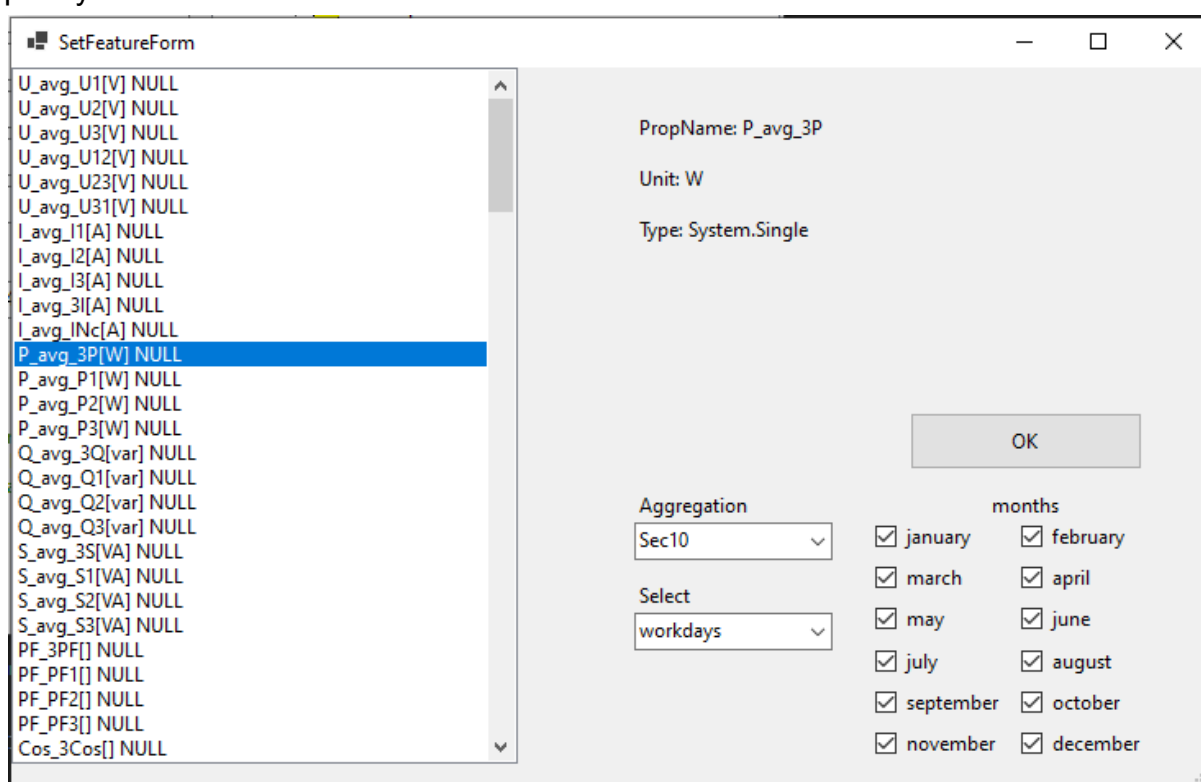
Navržený modul je součástí interního systému, který disponuje nástroji pro načtení dat. Tento návod neobsahuje postup pro načtení dat, ale zaměřuje se výhradně na práci s již načtenými daty v rámci modulu.

Po nahrání dat do systému se zobrazí úvodní okno s horním panelem tlačítek.



1. Výběr a filtrace atributu datasetu

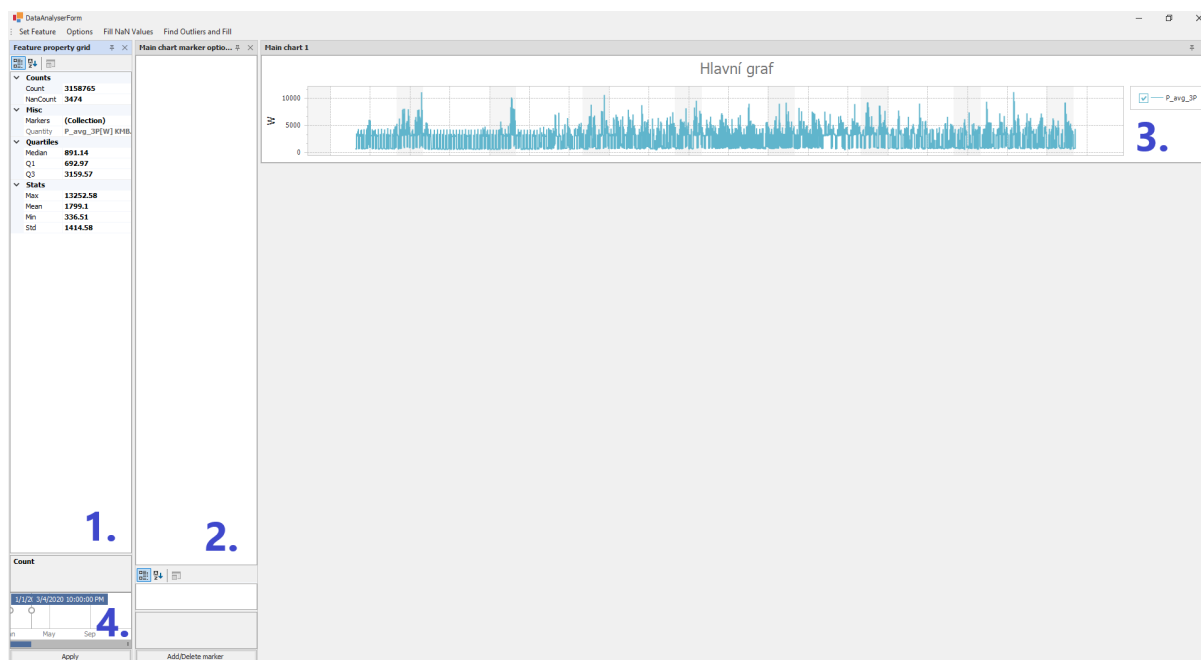
Stisknutím tlačítka “Set Feature” v horním panelu úvodního okna, se zobrazí okno pro výběr atributu.



- V levé části okna lze vybrat právě jeden atribut ze seznamu atributů nahraného datasetu.
- V poli “Aggregation” lze vybrat agregační interval. K dispozici jsou přesně definované intervaly. Je možné zvolit i možnost “None”, tedy bez agregačního intervalu.
- V poli “Select” lze filtrovat měření podle toho, zda se jedná o všední dny (workdays) nebo víkendy (weekends).
- Pomocí zaškrtnutých polí lze filtrovat měření podle měsíců.

Stisknutím tlačítka “OK” se potvrdí výběr a provede se selekce dat.

Po selekci dat se v úvodním okně zobrazí panely.

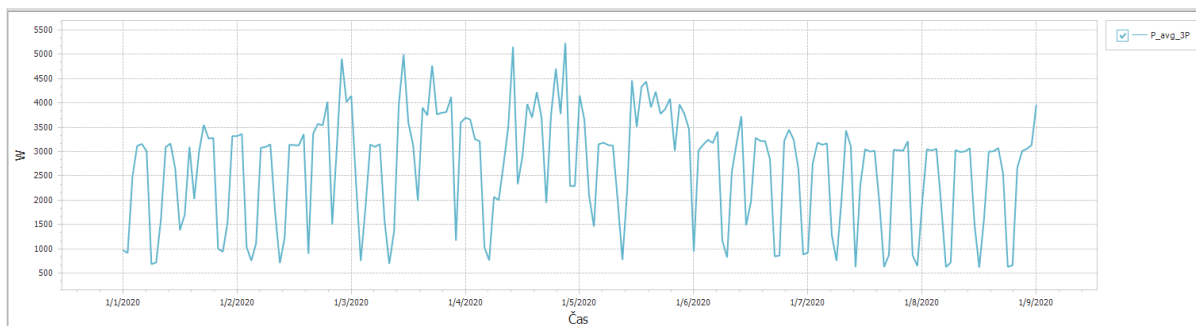


1. Panel obsahující tabulku popisných dat vybraného atributu.
 - “Count” (počet validních vzorků),
 - “NanCount” (počet chybějících vzorků),
 - “Markers” (seznam značek),
 - “Quantity” (název vybraného atributu),
 - “Median” (medián),
 - “Q1” (první kvartil),
 - “Q3” (třetí kvartil),
 - “Max” (maximální hodnota),
 - “Mean” (průměrná hodnota),
 - “Min” (minimální hodnota),
 - “Std” (směrodatná odchylka)
2. Panel obsahující seznam značek - viz [4. Značky](#).
3. Panel obsahující hlavní graf.
4. Panel obsahující časovou osu a tlačítko pro výběr časového intervalu.
 - Podržením levého tlačítka myši v poli časové osy lze vybrat časový interval.
 - Stisknutím pravého tlačítka myši v poli časové osy se nastaví týdenní časový interval.
 - Stisknutím tlačítka “Apply” se vybraný časový interval aplikuje a případně se projeví v aktivních grafech.

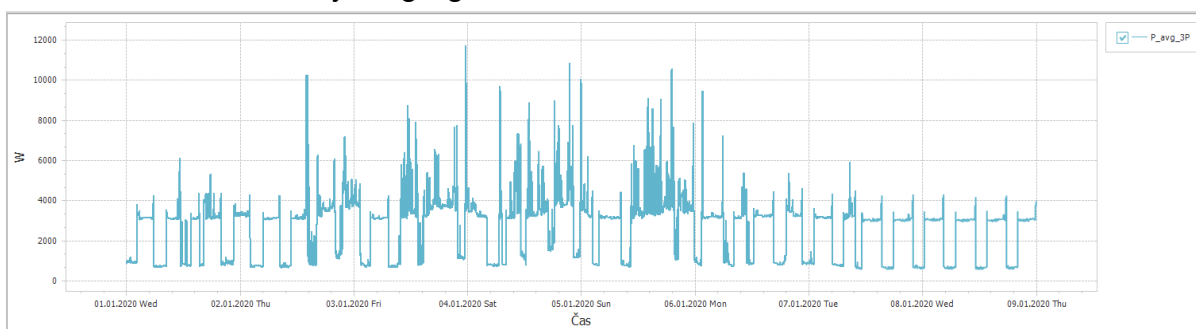
2. Vizualizace dat

Stisknutím tlačítka “Options” v horním panelu úvodního okna se zobrazí seznam tlačítek.

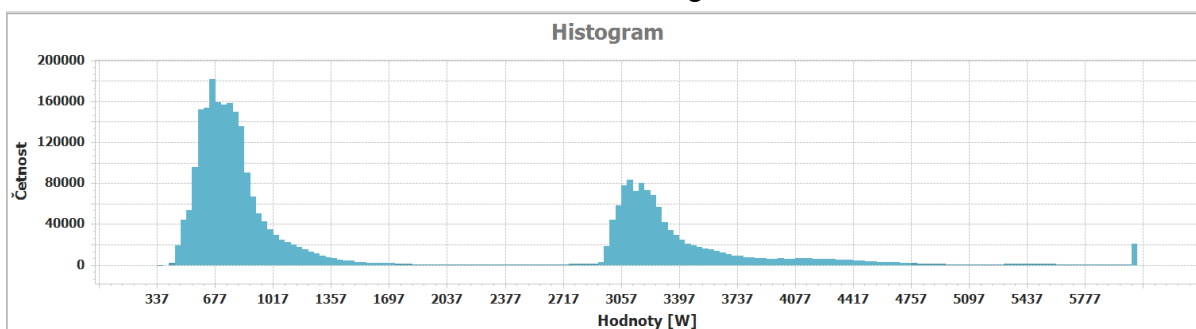
- “Line Graph”
 - Stisknutím tlačítka se zobrazí spojnicový graf s automatickou agregací dat, tj. v závislosti na úrovni přiblížení se mění agregační interval.



- “Line Graph (seconds)”
 - Stisknutím tlačítka se zobrazí spojnicový graf se statickým sekundovým agregačním intervalem.

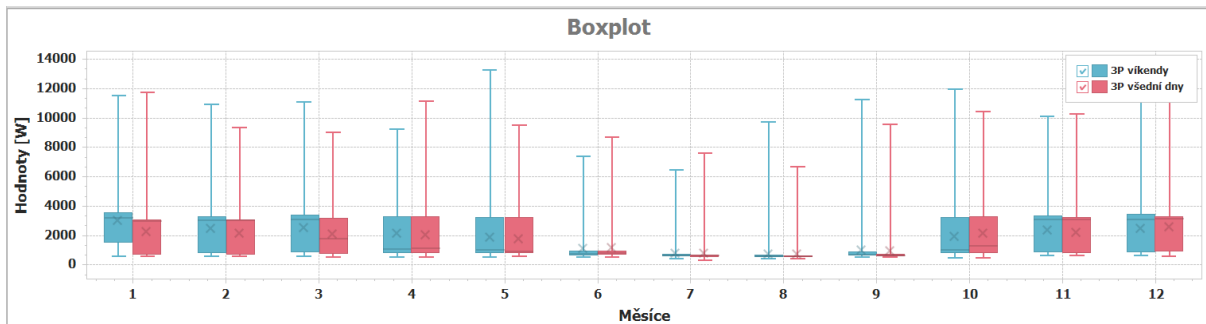


- “Histogram”
 - Stisknutím tlačítka se zobrazí histogram.

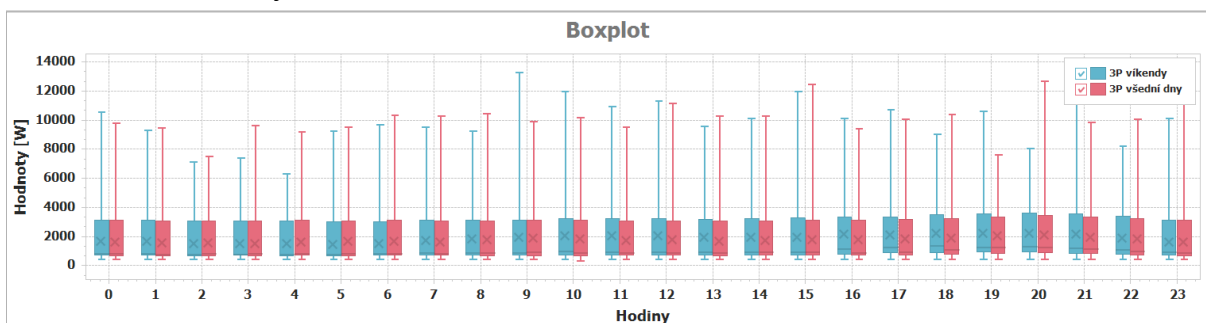


- “Markers” - viz [3. Metody pro analýzu dat](#).

- “Boxplot (month)”
 - Před použitím je nutné data očistit od chybějících hodnot. (viz [5. Chybějící a odlehlé hodnoty](#))
 - Stisknutím tlačítka se zobrazí krabicový graf s rozdělením dat podle měsíce měření.



- “Boxplot (hour)”
 - Před použitím je nutné data očistit od chybějících hodnot. (viz 5. Chybějící a odlehlé hodnoty)
 - Stisknutím tlačítka se zobrazí krabicový graf s rozdělením dat podle hodiny měření.

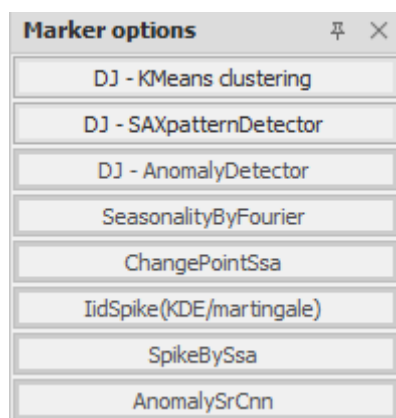


2.1. Ovládání grafů

- **Přiblížení**
 - Podržením klávesy “shift” a levého tlačítka myši v poli grafu lze přiblížit vybraný interval.
 - Podržením klávesy “shift” a stisknutím levého tlačítka myši v poli grafu lze přiblížit oblast kurzoru.
- **Oddálení**
 - Podržením klávesy “alt” a stisknutím levého tlačítka myši v poli grafu lze oddálit křivky.
- **Legenda**
 - Za pomoci zaškrtačkových tlačítek v legendě lze skrýt a zobrazit vybrané křivky.

3. Metody pro analýzu dat

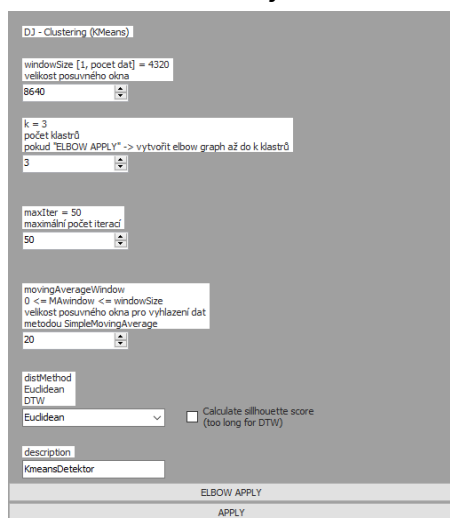
Stisknutím tlačítka “Options” v horním panelu úvodního okna se zobrazí seznam tlačítek. Stisknutím tlačítka “Markers” v seznamu se v pravém dolním rohu úvodního okna zobrazí panel se seznamem metod.



Stisknutím některého z tlačítek se zobrazí pod seznamem panel s nastavením parametrů pro vybranou metodu. Jednotlivé parametry jsou v panelu stručně popsány.

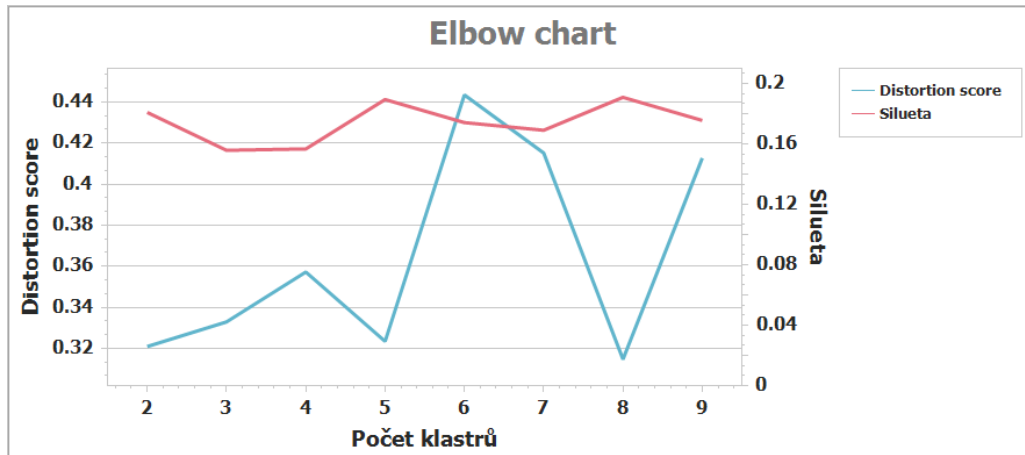
3.1. DJ - KMeans clustering

- Metoda shlukování metodou nejbližších středů

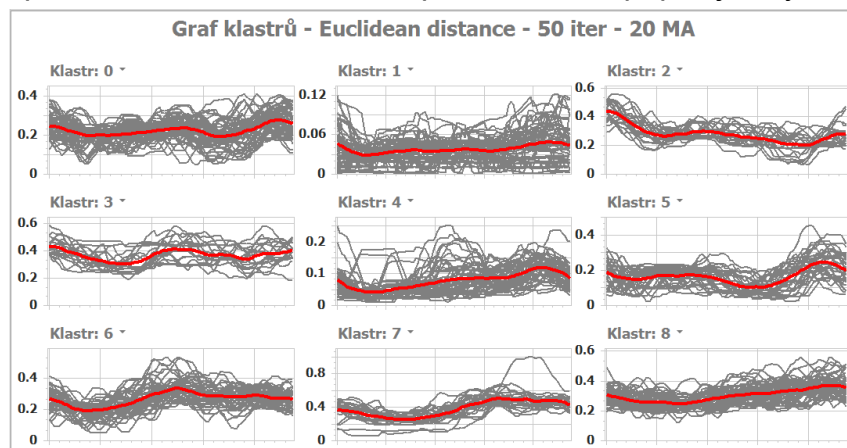


- Parametry:
 - “windowSize” (velikost posuvného okna, tj. kolik vzorků tvoří jeden vzorek v procesu klastrování),
 - “k” (výsledný počet klastrů),
 - “maxIter” (maximální počet iterací),
 - “movingAverageWindow” (velikost posuvného okna),
 - “distMethod” (vzdálenostní metrika, 2 možnosti - euklidovská metrika (Euclidean) a dynamic time warping (DTW)),
 - “Calculate silhouette score” (zda během zpracování metody počítat siluetu),
 - “description” (popis výsledné značky)

- Stisknutím tlačítka “ELBOW APPLY” se vytvoří Elbow chart a do panelu seznamu značek se přidá element popisující výsledek metody.



- Stisknutím tlačítka “APPLY” se vytvoří graf s nalezenými klastry a do panelu seznamu značek se přidá element popisující výsledek metody.



3.2. DJ - SAXpatternDetector

- Metoda symbolic aggregate approximation

DJ - SAX pattern detector

windowSize [1, počet dat] = 4320
velikost posuvného okna
8640

windowSplit [1, počet dat] = 4
na kolik částí se rozdělí okno
4

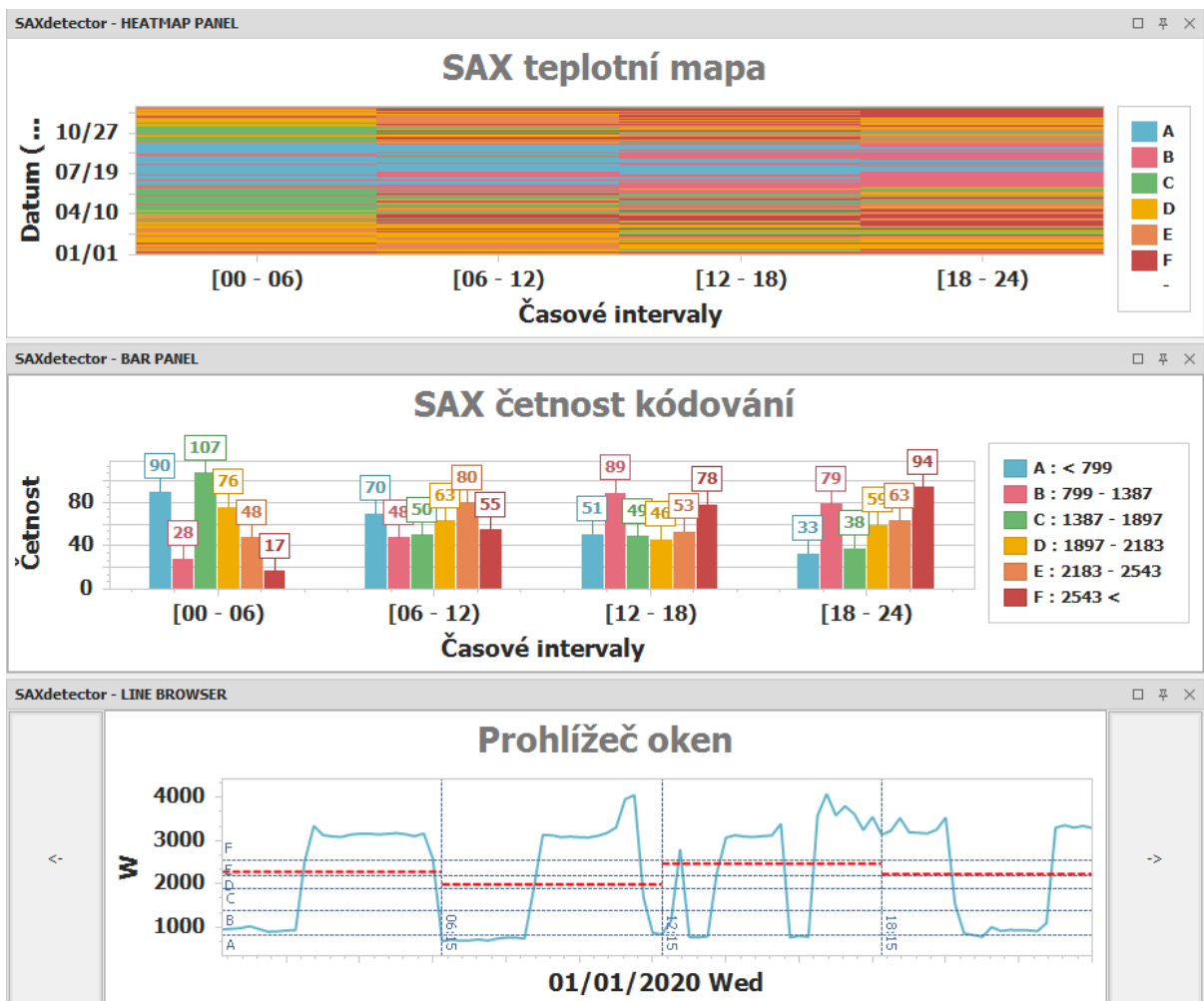
boundaries
pole hodnot vnitřních hranic
dolní a horní hranice již určeny
pokud == -1, x
- x hodnota definuje počet znaků
- hranice jsou vybrány podle kvantilů,
tak aby data byla rozdělena rovnoměrně
1000, 2000, 3000, 4000

threshold
== 0 -> nezobrazují anomálie
> 0 -> najdi data, která jsou o threshold vedle
0

description
SAXdetector ☒ Show panel

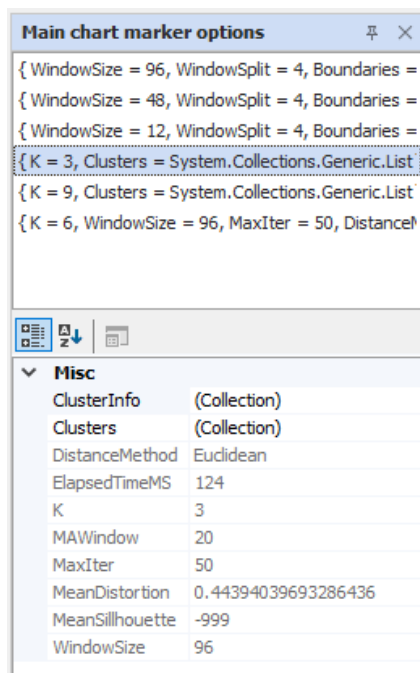
APPLY

- Parametry:
 - “windowSize” (velikost posuvného okna, tj. kolik vzorků tvoří jeden vzorek v metodě symbolic aggregate approximation),
 - “windowSplit” (na kolik vnitřních oken se rozdělí hlavní okno),
 - “boundaries” (pole hodnot vnitřních hranic),
 - “description” (popis výsledné značky),
 - “Show panel” (zda nechat zobrazit výsledné grafy)
- Stisknutím tlačítka “APPLY” se vytvoří 3 grafy (teplotní mapa, četnost kódování a prohlížeč oken) a do panelu seznamu značek se přidá element popisující výsledek metody.



4. Značky

Při aplikaci libovolné metody se její výsledky a statistiky uloží jako element v seznamu značek, který je k dispozici v panelu úvodního okna.



Výběrem libovolného elementu ze seznamu pomocí levého tlačítka myši se jeho položky zobrazí v tabulce pod seznamem.

5. Chybějící a odlehlé hodnoty

- Stisknutím tlačítka “Fill NaN Values” v horním panelu úvodního okna lze doplnit chybějící hodnoty.
- Stisknutím tlačítka “Find Outliers and Fill” v horním panelu úvodního okna lze detekovat odlehlé hodnoty a nahradit je validními hodnotami.