Spis treści

[Obserwacje wynikające z Eksploracyjnej Analizy Danych 2](#_Toc134290193)

[Kiedy przysyłanych jest najwięcej dokumentów? 3](#_Toc134290194)

[Duration X dzień X godzina 5](#_Toc134290195)

[Duration X dzień X tydzień 6](#_Toc134290196)

[Jakie są osiągane czasy przetwarzania? 11](#_Toc134290197)

[Eksploracyjna Analiza Danych 13](#_Toc134290198)

[**Zbiór danych** - typy danych 13](#_Toc134290199)

[**Zbiór danych** - generalne statystyki 13](#_Toc134290200)

[**Zbiór danych** - ostrzeżenia 14](#_Toc134290201)

[**Zmienna - Queue** - charakterystyki zmiennej 15](#_Toc134290202)

[**Zmienna - InputPageCount** - charakterystyki zmiennej 16](#_Toc134290203)

[**Zmienna - MessagesInQueue** - charakterystyki zmiennej 17](#_Toc134290204)

[**Zmienna - Retries** - charakterystyki zmiennej 20](#_Toc134290205)

[**Zmienna - ProcessingDuration\_sec** - charakterystyki zmiennej 22](#_Toc134290206)

[**Zmienna - Duration\_sec** - charakterystyki zmiennej 24](#_Toc134290207)

[**Zmienna - StartDateTimeUtc** - charakterystyki zmiennej 26](#_Toc134290208)

[**Tablice korelacji** – korelacja między zmiennymi 27](#_Toc134290209)

[**Tablice korelacji** – korelacja między brakami danych 27](#_Toc134290210)

# Obserwacje wynikające z Eksploracyjnej Analizy Danych

* Dla celów wnioskowania na temat statystyk całego Zbioru Danych odrzucam zdarzenia, gdy pole *ID\_Document* jest wartością pustą. Innymi słowy: w celu wyliczenia Statystyk Globalnych odrzucam te zdarzenia z dalszej analizy. Dokumentów bez przypisanego Potoku w Zbiorze Danych jest 1.2% . Nie jest mi znana interpretacja biznesowa zdarzeń, gdy *ID\_Document* przyjmuje wartość pustą.
* Analizując wszystkie Potoki razem okazuje się, że nie wszystkie potoki były funkcjonalne każdego dnia (przykład: potok\_3-2). Być może wynikało to z awarii na Środowisku Produkcyjnym lub otrzymany Zbiór Danych zawiera informacje o zdarzeniach ze Środowiska Deweloperskiego, w tym Zdarzenia związane z okresami Testowymi. Ze względu na powyższe zaleca się analizę każdego z Potoków osobno. Zaleca się wyciąganie wniosków oraz tworzenie rekomendacji w oparciu o statyki dla każdego z Potoków osobno.
* Wszystkie Zmienne numeryczne opisujące Potoki: *Duration*, *ProcessingDuration*, *InputPageCount*, *Retries*, ze względu na swój charakter biznesowy – procesowanie faktur w oparciu o rozwiązanie OCR (ang. *optical character recognition*), charakteryzują się rozkładami prawo skośnymi. Najczęściej Klienci procesują pojedyncze faktury.
* Historia zdarzeń potok\_2-1 (1193) oraz potok\_2-2 (497) dotyczy wyłącznie jednego dnia – tzn. poniedziałku 8 tygodnia 2023 roku. Wydaje się niezasadnym wnioskowanie na podstawie danego potoku w oparciu o historię zdarzeń z wyłącznie jednego dnia uwzględnienie. Zdarzenia związane z powyższymi potokami nie zostały uwzględnione w dalszych analizach.

# Kiedy przysyłanych jest najwięcej dokumentów?

Rozkład zmiennej będącej flagą czasu zdarzenia prezentuje się następująco:  
  
Obraz zawierający tekst, zrzut ekranu, Czcionka, diagram

Opis wygenerowany automatycznie

Na podstawie powyższego histogramu ciężko jednoznacznie stwierdzić kiedy i dlaczego procesowany jest więcej dokumentów.

Z dużą dozą pewności założyć można, że więcej dokumentów procesowanych jest w tzw. godzinach biznesowych, tzw. dni pracujących. Czyli mniej więcej od 8:00 rano do 17:00 popołudniu. Warto to sprawdzić.

Faktycznie histogram ze skumulowanej liczebności zdarzeń dla każdego z dni tygodnia potwierdzaja intuicyjną tezę, że w łikendy procesowanych jest mniej faktur.

Obraz zawierający zrzut ekranu, diagram, tekst, Prostokąt

Opis wygenerowany automatycznie

Niemniej w łikendy wciąż procesuje się znaczne liczby dokumentów. Co potwierdzają poniższe skumulowane liczebności w podziale na tygodnie:

[Histogram skumulowanych liczebności zdarzeń w podziale na dni oraz tygodnie](https://github.com/KonuTech/yet-another-analysis/blob/main/03_hisotgram_stacked_days_x_weeks.png)

Warto zauważyć, że na przestrzeni dwóch miesięcy zmniejszyła się liczba zdarzeń związanych z potok\_1-1 na rzecz zdarzeń związanych z potok\_2-1.

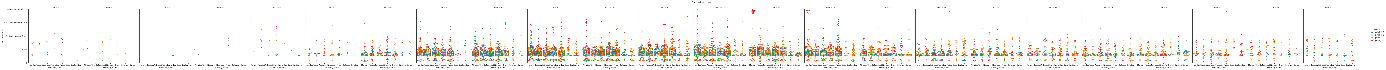
Większy popyt na usługi procesowania faktur w czasie tzw. godzin biznesowych powoduje, że część faktur trafia w tzw. kolejkę. Tym samym w dalszej części analiz do wymiarów czasu: godzin, dnia, tygodnia, miesięcy dodam zmienne *Duration* oraz *ProcessingDuration*.

Tzw. wykresy rojowe pozwalają ocenić alokację zdarzeń w podzielę na kategorie.

## Duration X dzień X godzina

Poniższy wykres przedstawia próbkę 10% zdarzeń (około 100 tys. zdarzeń) w podziale na dni tygodnia i godziny. Wykres pozwala na wstępną ocenę popytu na Usługi z uwzględnieniem wymiaru czasu.

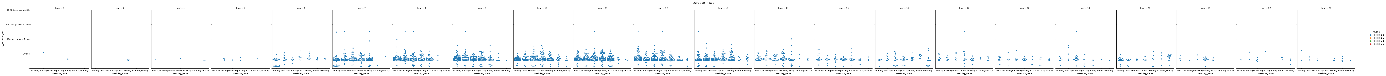
**Wszystkie potoki:**



Większy wykres:  
[Wykres roju dla wszystkich potoków – Duration w sek. dla dla godziny i tygodnia](https://raw.githubusercontent.com/KonuTech/yet-another-analysis/main/01_swarm_duration_x_days_x_hours.png)

Z punktu widzenia każdego z potoków z osobna, wykresy rojowe przedstawiają podobny popyt na usługi w interakcji pomiędzy dniami tygodnia a godzinami.

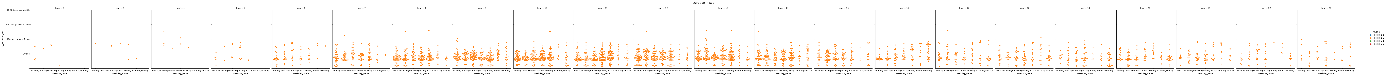
**potok\_1-1:**



Większy wykres:

[Wykres roju potok\_1-1 – Duration w sek. dla godziny i tygodnia](https://raw.githubusercontent.com/KonuTech/yet-another-analysis/main/01_swarm_potok_1_1_duration_x_days_x_hours.png)

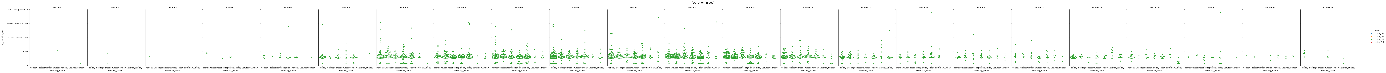
**potok\_1-2:**



Większy wykres:

[Wykres roju potok\_1-2 – Duration w sek. dla godziny i tygodnia](https://raw.githubusercontent.com/KonuTech/yet-another-analysis/main/01_swarm_potok_1_2_duration_x_days_x_hours.png)

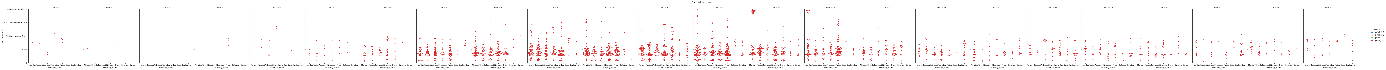
**potok\_3-1:**



Większy wykres:

[Wykres roju potok\_3-1 – Duration w sek. dla godziny i tygodnia](https://raw.githubusercontent.com/KonuTech/yet-another-analysis/main/01_swarm_potok_3_1_duration_x_days_x_hours.png)

**potok\_3-2:**



Większy wykres:

[Wykres roju potok\_3-2 – Duration w sek. dla godziny i tygodnia](https://github.com/KonuTech/yet-another-analysis/blob/main/01_swarm_potok_3_2_duration_x_days_x_hours.png)

Spostrzeżenia:

Z perspektywy wszystkich Potoków, najwięcej dokumentów procesowanych jest:

* między godziną 7:00 a 15:00
* wciąż sporo faktur procesowanych jest między 5:00 a 6:00 oraz między 15:00 a 20:00
* znacznie mniej faktur procesowanych jest między 20:00 a 24:00
* istotnie niewiele faktur procesowanych jest między 24:00 a 6:00.

## Duration X dzień X tydzień

Poniższy wykres przedstawia 10% zdarzeń (około 100 tys. zdarzeń) w podziale na dni tygodnia i tygodnie otrzymanego Zbioru danych. Wykresy pozwalają na ocenę obciążenia Usług w wymiarze czasu.

[Wykres roju dla wszystkich potoków – Duration w sek. dla dnia i tygodnia](https://github.com/KonuTech/yet-another-analysis/blob/main/02_swarm_duration_x_days_x_weeks.png)

Wykresy rojowe dla każdego z Potoków przedstawiają podobne zaangażowanie zdarzeń w interakcji pomiędzy dniem tygodnia a godziną.

**potok\_1-1:**

Obraz zawierający tekst, linia, Wykres, Czcionka

Opis wygenerowany automatycznie

Większy wykres:

[Wykres roju potok\_1-1 – Duration w sek. dla dnia i tygodnia](https://github.com/KonuTech/yet-another-analysis/blob/main/02_swarm_potok_1_1_duration_x_days_x_weeks.png)

**potok\_1-2:**

Obraz zawierający tekst, linia, Wykres, Czcionka

Opis wygenerowany automatycznie

Większy wykres:

[Wykres roju potok\_1-2 – Duration w sek. dla dnia i tygodnia](https://github.com/KonuTech/yet-another-analysis/blob/main/02_swarm_potok_1_2_duration_x_days_x_weeks.png)

**potok\_3-1:**

Obraz zawierający tekst, zrzut ekranu, linia, Wykres

Opis wygenerowany automatycznie

Większy wykres:

[Wykres roju potok\_3-1 – Duration w sek. dla dnia i tygodnia](https://github.com/KonuTech/yet-another-analysis/blob/main/02_swarm_potok_2_1_duration_x_days_x_weeks.png)

**potok\_3-2:**

Obraz zawierający tekst, linia, Czcionka, Wykres

Opis wygenerowany automatycznie

Większy wykres:

[Wykres roju potok\_3-2 – Duration w sek. dla dnia i tygodnia](https://github.com/KonuTech/yet-another-analysis/blob/main/02_swarm_potok_2_2_duration_x_days_x_weeks.png)

Spostrzeżenia:

* We wtorek i środę 8 tygodnia 2023 r. wszystkie potoki były wyłączone lub doznały awarii.
* Między tygodniem 5 a 11 2023 r. jedynie potok\_1-1 oraz potok\_1-2 reprezentowały stabilną historię zdarzeń.
* Być może w tygodniu 5 i 6 potok\_1-1 doświadczył awarii. *Duration* dla potoku\_1-1 było w tygodniu 5 wyraźnie wyższe aniżeli w tygodniach 7-11, kiedy to potok\_1-1 funkcjonował stabilnie. W poniedziałek w tygodniu 6 wiele procesów potoku\_1-1 charakteryzowało *Duration* nawet na poziomie 3h.
* Warto zauważyć, że:
  + ze względu na występowanie nietypowych zdarzeń średnia *Duration* dla całego zbioru danych to 25 sekundy.
  + natomiast historyczna mediana *Duration* wynosi poniżej 4 sekund
  + z kolei pierwszy kwartyl wynosi poniżej 3 sekund
  + a 3 kwartyl wynosi 7 sekund.

Jako, że każde unikatowe ID\_Document może być przetwarzane wiele razy, to do oceny kiedy przesyłanych jest najwięcej dokumentów użyję agregaty ID\_Process.

Na pytanie kiedy przysyłanych jest najwięcej dokumentów mogą odpowiedzieć nam m. in. trzy miary: średnia liczba procesów, średnie *Duration procesów* oraz średnie *MessagesInQueue* procesu - dla każdej godziny. Uwzględniając liczbę miesięcy i dni.

Według poniższego wykresu liniowego, najwięcej dokumentów przesyłanych jest między 6:00 a 15:00.

Obraz zawierający tekst, diagram, Wykres, linia

Opis wygenerowany automatycznie

Wszystkie potoki:Obraz zawierający tekst, zrzut ekranu, Wykres, diagram

Opis wygenerowany automatycznie

Wszystkie potoki z wyłączeniem potok\_3-1

Obraz zawierający tekst, diagram, Wykres, linia

Opis wygenerowany automatycznie

Wykres *MessagesInQueue* w

Obraz zawierający tekst, Wykres, diagram, zrzut ekranu

Opis wygenerowany automatycznie

# Jakie są osiągane czasy przetwarzania? (zmienna ProcesingDuration)

W każdym z potoków wystąpiły zdarzenia nazywane zdarzeniami„ekstremalnymi”, „odstającymi”.

Czasy przetwarzania były rozmaite dla wszystkich potoków.

Obraz zawierający tekst, zrzut ekranu, diagram, Wykres

Opis wygenerowany automatycznie

Zdarzenia odstające występowały każdego dnia tygodnia

Obraz zawierający tekst, zrzut ekranu, diagram, Wykres

Opis wygenerowany automatycznie

Zdarzenia odstające występowały zarówno w typowych godzinach biurowych jak i poza typowymi godzinami biurowymi, czy w łikendy.

potok\_1-1 (próbka 10%)

Obraz zawierający statek wodny, panorama

Opis wygenerowany automatycznie przy średnim poziomie pewności

potok\_1-2 (próbka 10%)

Obraz zawierający statek, statek wodny, panorama

Opis wygenerowany automatycznie

potok\_3-1 (próbka 10%)

Obraz zawierający statek, statek wodny

Opis wygenerowany automatycznie

potok\_3-2 (próbka 10%)

Obraz zawierający panorama, statek wodny

Opis wygenerowany automatycznie

Osiągane czasy przetwarzania dla wszystkich potoków wyglądają podobnie. Wydaje się, że zarówno w przypadku potok\_1-1 jak i potok\_3-1 zdarzenie odstające występują nieco rzadziej aniżeli w przypadku potok\_1-2, potok\_3-1. Być może istnieje w tym zjawisku interpretacja biznesowo-techniczna, która nie jest mi w tej chwili znana.

Drążenie danych po tygodniu potwierdza, że na przestrzeni obu miesięcy, zdarzenia odstające nie były przypadkiem.

Obraz zawierający tekst, diagram, linia, zrzut ekranu

Opis wygenerowany automatycznie

Obraz zawierający tekst, diagram, zrzut ekranu, linia

Opis wygenerowany automatycznie

Obraz zawierający diagram, linia, tekst, zrzut ekranu

Opis wygenerowany automatycznie

Obraz zawierający diagram, linia, tekst, Wykres

Opis wygenerowany automatycznie

Obraz zawierający tekst, linia, diagram, Wykres

Opis wygenerowany automatycznie

Obraz zawierający diagram, linia, tekst, Wykres

Opis wygenerowany automatycznie

Obraz zawierający diagram, tekst, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający diagram, tekst, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający tekst, diagram, linia, zrzut ekranu

Opis wygenerowany automatycznie

We wszystkich tygodniach doświadczano zdarzeń odstających

Obraz zawierający tekst, dokument, Czcionka, papier

Opis wygenerowany automatycznie

# Jakie są osiągane czasy oczekiwania na wynik? (zmienna Duration)

W każdym z potoków wystąpiły zdarzenia nazywane zdarzeniami„ekstremalnymi”, „odstającymi”.

Czasy przetwarzania były rozmaite dla wszystkich potoków.

Obraz zawierający tekst, zrzut ekranu, diagram, Wykres

Opis wygenerowany automatycznie

Zdarzenia odstające występowały każdego dnia tygodnia

Obraz zawierający tekst, zrzut ekranu, diagram, Wykres

Opis wygenerowany automatycznie

Zdarzenia odstające występowały zarówno w typowych godzinach biurowych jak i poza typowymi godzinami biurowymi, czy w łikendy.

potok\_1-1 (próbka 10%)

Obraz zawierający statek wodny

Opis wygenerowany automatycznie przy średnim poziomie pewności

potok\_1-2 (próbka 10%)

Obraz zawierający statek, statek wodny, panorama

Opis wygenerowany automatycznie

potok\_3-1 (próbka 10%)

Obraz zawierający diagram, linia

Opis wygenerowany automatycznie

potok\_3-2 (próbka 10%)

Obraz zawierający panorama

Opis wygenerowany automatycznie

Osiągane czasy przetwarzania dla wszystkich potoków wyglądają podobnie. Wydaje się, że zarówno w przypadku potok\_1-1 jak i potok\_3-1 zdarzenie odstające występują nieco rzadziej aniżeli w przypadku potok\_1-2, potok\_3-1. Być może istnieje w tym zjawisku interpretacja biznesowo-techniczna, która nie jest mi w tej chwili znana.

Drążenie danych po tygodniu potwierdza, że na przestrzeni obu miesięcy, zdarzenia odstające nie były przypadkiem.

Obraz zawierający diagram, linia, Wykres, tekst

Opis wygenerowany automatycznie

Obraz zawierający diagram, zrzut ekranu, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający diagram, zrzut ekranu, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający diagram, zrzut ekranu, tekst, linia

Opis wygenerowany automatycznie

Obraz zawierający diagram, zrzut ekranu, tekst, linia

Opis wygenerowany automatycznie

Obraz zawierający diagram, tekst, Wykres, linia

Opis wygenerowany automatycznie

Obraz zawierający diagram, tekst, zrzut ekranu, linia

Opis wygenerowany automatycznie

Obraz zawierający diagram, tekst, linia, zrzut ekranu

Opis wygenerowany automatycznie

Obraz zawierający tekst, diagram, linia, zrzut ekranu

Opis wygenerowany automatycznie

We wszystkich tygodniach doświadczano zdarzeń odstających

Obraz zawierający tekst, Czcionka, dokument, pismo odręczne

Opis wygenerowany automatycznie

# Eksploracyjna Analiza Danych

## **Zbiór danych** - typy danych

|  |  |
| --- | --- |
| Numeryczna | 7 |
| Kategoria | 1 |
| Data-czas | 1 |

## **Zbiór danych** - generalne statystyki

|  |  |
| --- | --- |
| Liczba zmiennych | 9 |
| Liczba obserwacji | 1121190 |
| Braki danych | 139479 |
| Braki danych (%) | 1.4% |
| Zduplikowane wiersze | 0 |
| Zduplikowane wiersze (%) | 0.0% |

## **Zbiór danych** - ostrzeżenia

|  |  |
| --- | --- |
| [ID\_Process](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_8499889631538707095) jest silnie skorelowany z ID\_Document | **High correlation** |
| [ID\_Document](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_-5726450493592489171) jest silnie skorelowany z ID\_Process | **High correlation** |
| [ProcessingDuration\_sec](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_-2639043556186121784) jest silnie skorelowany z Duration\_sec | **High correlation** |
| [Duration\_sec](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_4501912086993125054) jest silnie skorelowany z ProcessingDuration\_sec | **High correlation** |
| [Queue](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_-4899536447608545937) ma 13366 (1.2%) braków danych | **Missing** |
| [ID\_Document](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_-5726450493592489171) ma 13366 (1.2%) braków danych | **Missing** |
| [InputPageCount](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_8935948530309897777) ma 35784 (3.2%) braków danych | **Missing** |
| [ProcessingDuration\_sec](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_-2639043556186121784) ma 35784 (3.2%) braków danych | **Missing** |
| [Duration\_sec](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_4501912086993125054) ma 41179 (3.7%) braków danych | **Missing** |
| [InputPageCount](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_8935948530309897777) ma silnie skośny rozkład (γ1 = 147.8433905) | **Skewed** |
| [MessagesInQueue](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_39819409819331012) ma silnie skośny rozkład (γ1 = 26.18552423) | **Skewed** |
| [Retries](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_-1867881654231547758) ma silnie skośny rozkład (γ1 = 35.33890113) | **Skewed** |
| [ProcessingDuration\_sec](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_-2639043556186121784) ma silnie skośny rozkład (γ1 = 39.97939455) | **Skewed** |
| [Duration\_sec](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_4501912086993125054) ma silnie skośny rozkład (γ1 = 35.78904802) | **Skewed** |
| [ID\_Process](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_8499889631538707095) ma unikatowe wartości | **Unique** |
| [MessagesInQueue](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_39819409819331012) ma 247117 (22.0%) zer | **Zeros** |
| [Retries](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_-1867881654231547758) ma 1116661 (99.6%) zer | **Zeros** |

## **Zmienna -** [**Queue**](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_-4899536447608545937) - charakterystyki zmiennej

|  |  |  |
| --- | --- | --- |
| Wartości unikatowe | | 6 |
| Wartości unikatowe (%) | | < 0.1% |
| Braki danych | | 13366 |
| Braki danych (%) | | 1.2% |
| **Wartość** | **Liczebność** | **Częstotliwość (%)** |
| potok\_1-2 | 333292 | 29.7% |
| potok\_1-1 | 271952 | 24.3% |
| potok\_3-1 | 255158 | 22.8% |
| potok\_3-2 | 245732 | 21.9% |
| potok\_2-1 | 1193 | 0.1% |
| potok\_2-2 | 497 | < 0.1% |
| (Braki danych) | 13366 | 1.2% |

Obraz zawierający tekst, zrzut ekranu, Czcionka, design

Opis wygenerowany automatycznie

## [**Zmienna - InputPageCount**](http://localhost:63342/yet-another-analysis/outputs/report_timeseries.html?_ijt=tapd920bd1e4gb0nrgmalqbhcu&_ij_reload=RELOAD_ON_SAVE#pp_var_8935948530309897777) - charakterystyki zmiennej

|  |  |
| --- | --- |
| Wartości unikatowe | 109 |
| Wartości unikatowe **(%)** | < 0.1% |
| Braki danych | 35784 |
| Braki danych **(%)** | 3.2% |
| **Średnia** | 1.2139282 |
| **Minimum** | 1 |
| **Maksimum** | 736 |
| **Liczba zer** | 0 |
| **Liczba zer (%)** | 0.0% |
| **Liczba wartości negatywnych** | 0 |
| **Liczba wartości negatywnych (%)** | 0.0% |

Statystyki kwantylowe

|  |  |
| --- | --- |
| **Minimum** | 1 |
| **5-th percentyl** | 1 |
| **Q1** | 1 |
| **mediana** | 1 |
| **Q3** | 1 |
| **95-th percentyl** | 2 |
| **Maksimum** | 736 |
| **Rozstęp** | 735 |
| **Rozstęp miedzykwartylowy (IQR)** | 0 |

Statystyki opisowe

|  |  |
| --- | --- |
| **Odchylenie standardowe** | 2.1322051 |
| **Współczynnik wariancji (CV)** | 1.7564507 |
| **Kurtoza** | 34862.734 |
| **Średnia** | 1.2139282 |
| **Median Absolute Deviation (MAD)** | 0 |
| **Skośność** | 147.84339 |
| **Sum** | 1317605 |
| **Wariancja** | 4.5462985 |

Wartości najczęstsze - 10 wartości minimalnych

| **Wartość** | **Liczebność** | **Częstotliwość (%)** |
| --- | --- | --- |
| 1 | 980867 | 87.5% |
| 2 | 65159 | 5.8% |
| 3 | 18571 | 1.7% |
| 4 | 8845 | 0.8% |
| 5 | 3131 | 0.3% |
| 6 | 3113 | 0.3% |
| 7 | 1221 | 0.1% |
| 8 | 1103 | 0.1% |
| 10 | 568 | 0.1% |
| 9 | 564 | 0.1% |
| Pozostałe wartości | 2264 | 0.2% |
| (Braki danych) | 35784 | 3.2% |

Wartości ekstremalne - 10 wartości maksymalnych

|  |
| --- |
|  |
| **Wartość** | **Liczebność** | **Częstotliwość** (%) |
| 736 | 1 | < 0.1% |
| 579 | 1 | < 0.1% |
| 530 | 1 | < 0.1% |
| 514 | 1 | < 0.1% |
| 459 | 2 | < 0.1% |
| 394 | 1 | < 0.1% |
| 365 | 1 | < 0.1% |
| 355 | 2 | < 0.1% |
| 350 | 1 | < 0.1% |
| 342 | 1 | < 0.1% |

## **Zmienna - MessagesInQueue** - charakterystyki zmiennej

|  |  |
| --- | --- |
| Wartości unikatowe | 1670 |
| Wartości unikatowe **(%)** | 0.1% |
| Braki danych | 0 |
| Braki danych **(%)** | 0.0% |
| **Średnia** | 9.009044 |
| **Minimum** | 0 |
| **Maksimum** | 2018 |
| **Liczba zer** | 247117 |
| **Liczba zer (%)** | 22.0% |
| **Liczba wartości negatywnych** | 0 |
| **Liczba wartości negatywnych (%)** | 0.0% |

Statystyki kwantylowe

|  |  |
| --- | --- |
| **Minimum** | 0 |
| **5-th percentyl** | 0 |
| **Q1** | 1 |
| **mediana** | 3 |
| **Q3** | 7 |
| **95-th percentyl** | 30 |
| **Maksimum** | 2018 |
| **Rozstęp** | 2018 |
| **Rozstęp miedzykwartylowy (IQR)** | 6 |

Statystyki opisowe

|  |  |
| --- | --- |
| **Odchylenie standardowe** | 52.464583 |
| **Współczynnik wariancji (CV)** | 5.8235461 |
| **Kurtoza** | 793.10816 |
| **Średnia** | 9.009044 |
| **Median Absolute Deviation (MAD)** | 3 |
| **Skośność** | 26.185524 |
| **Sum** | 10100850 |
| **Wariancja** | 2752.5324 |

Wartości najczęstsze - 10 wartości minimalnych

| **Wartość** | **Liczebność** | **Częstotliwość (%)** |
| --- | --- | --- |
| 0 | 247117 | 22.0% |
| 1 | 159603 | 14.2% |
| 2 | 139394 | 12.4% |
| 3 | 109755 | 9.8% |
| 4 | 81661 | 7.3% |
| 5 | 57776 | 5.2% |
| 6 | 42693 | 3.8% |
| 7 | 32380 | 2.9% |
| 8 | 26229 | 2.3% |
| 9 | 21117 | 1.9% |

Wartości ekstremalne - 10 wartości maksymalnych

|  |
| --- |
|  |
| **Wartość** | **Liczebność** | **Częstotliwość** (%) |
| 2018 | 1 | < 0.1% |
| 2017 | 1 | < 0.1% |
| 2016 | 1 | < 0.1% |
| 2015 | 1 | < 0.1% |
| 2011 | 4 | < 0.1% |
| 2010 | 1 | < 0.1% |
| 2007 | 3 | < 0.1% |
| 2006 | 1 | < 0.1% |
| 2005 | 1 | < 0.1% |
| 2003 | 1 | < 0.1% |

## **Zmienna - Retries** - charakterystyki zmiennej

|  |  |
| --- | --- |
| Wartości unikatowe | 10 |
| Wartości unikatowe **(%)** | < 0.1% |
| Braki danych | 0 |
| Braki danych **(%)** | 0.0% |
| **Średnia** | 0.0048956912 |
| **Minimum** | 0 |
| **Maksimum** | 9 |
| **Liczba zer** | 1116661 |
| **Liczba zer (%)** | 99.6% |
| **Liczba wartości negatywnych** | 0 |
| **Liczba wartości negatywnych (%)** | 0.0% |

Statystyki kwantylowe

|  |  |
| --- | --- |
| **Minimum** | 0 |
| **5-th percentyl** | 0 |
| **Q1** | 0 |
| **mediana** | 0 |
| **Q3** | 0 |
| **95-th percentyl** | 0 |
| **Maksimum** | 9 |
| **Rozstęp** | 9 |
| **Rozstęp miedzykwartylowy (IQR)** | 0 |

Statystyki opisowe

|  |  |
| --- | --- |
| **Odchylenie standardowe** | 0.090071803 |
| **Współczynnik wariancji (CV)** | 18.398179 |
| **Kurtoza** | 2320.6904 |
| **Średnia** | 0.0048956912 |
| **Median Absolute Deviation (MAD)** | 0 |
| **Skośność** | 35.338901 |
| **Sum** | 5489 |
| **Wariancja** | 0.0081129297 |

Wartości najczęstsze

| **Wartość** | **Liczebność** | **Częstotliwość (%)** |
| --- | --- | --- |
| 0 | 1116661 | 99.6% |
| 1 | 3912 | 0.3% |
| 2 | 451 | < 0.1% |
| 3 | 116 | < 0.1% |
|  | 19 | < 0.1% |
| 4 | 12 | < 0.1% |
| 5 | 12 | < 0.1% |
| 7 | 4 | < 0.1% |
| 6 | 2 | < 0.1% |
| 8 | 1 | < 0.1% |

## **Zmienna - ProcessingDuration\_sec** - charakterystyki zmiennej

|  |  |
| --- | --- |
| Wartości unikatowe | 27357 |
| Wartości unikatowe **(%)** | 2.5% |
| Braki danych | 35784 |
| Braki danych **(%)** | 3.2% |
| **Średnia** | 4.1875326 |
| **Minimum** | 0.034 |
| **Maksimum** | 1422.42 |
| **Liczba zer** | 0 |
| **Liczba zer (%)** | 0.0% |
| **Liczba wartości negatywnych** | 0 |
| **Liczba wartości negatywnych (%)** | 0.0% |

Statystyki kwantylowe

|  |  |
| --- | --- |
| **Minimum** | 0.034 |
| **5-th percentyl** | 0.103 |
| **Q1** | 2.264 |
| **mediana** | 2.973 |
| **Q3** | 4.841 |
| **95-th percentyl** | 11.352 |
| **Maksimum** | 1422.42 |
| **Rozstęp** | 1422.386 |
| **Rozstęp miedzykwartylowy (IQR)** | 2.577 |

Statystyki opisowe

|  |  |
| --- | --- |
| **Odchylenie standardowe** | 5.1955218 |
| **Współczynnik wariancji (CV)** | 1.240712 |
| **Kurtoza** | 7004.0239 |
| **Średnia** | 4.1875326 |
| **Median Absolute Deviation (MAD)** | 0.935 |
| **Skośność** | 39.979395 |
| **Sum** | 4545173 |
| **Wariancja** | 26.993447 |

Wartości najczęstsze

| **Wartość** | **Liczebność** | **Częstotliwość (%)** |
| --- | --- | --- |
| 0.079 | 1835 | 0.2% |
| 0.078 | 1806 | 0.2% |
| 0.077 | 1797 | 0.2% |
| 0.084 | 1747 | 0.2% |
| 0.082 | 1731 | 0.2% |
| 0.08 | 1684 | 0.2% |
| 0.076 | 1667 | 0.1% |
| 0.074 | 1666 | 0.1% |
| Pozostałe wartości | 1068199 | 95.3% |
| (Braki danych) | 35784 | 3.2% |

Wartości ekstremalne - 10 wartości maksymalnych

|  |
| --- |
|  |
| **Wartość** | **Liczebność** | **Częstotliwość** (%) |
| 1422.42 | 1 | < 0.1% |
| 825.557 | 1 | < 0.1% |
| 764.678 | 1 | < 0.1% |
| 620.435 | 1 | < 0.1% |
| 606.298 | 4 | < 0.1% |
| 543.082 | 1 | < 0.1% |
| 522.68 | 3 | < 0.1% |
| 415.748 | 1 | < 0.1% |
| 408.209 | 1 | < 0.1% |
| 380.062 | 1 | < 0.1% |

## **Zmienna - Duration\_sec** - charakterystyki zmiennej

|  |  |
| --- | --- |
| Wartości unikatowe | 71302 |
| Wartości unikatowe **(%)** | 6.6% |
| Braki danych | 41179 |
| Braki danych **(%)** | 3.7% |
| **Średnia** | 25.350377 |
| **Minimum** | 0.271 |
| **Maksimum** | 45199.914 |
| **Liczba zer** | 0 |
| **Liczba zer (%)** | 0.0% |
| **Liczba wartości negatywnych** | 0 |
| **Liczba wartości negatywnych (%)** | 0.0% |

Statystyki kwantylowe

|  |  |
| --- | --- |
| **Minimum** | 0.271 |
| **5-th percentyl** | 0.548 |
| **Q1** | 2.714 |
| **mediana** | 3.707 |
| **Q3** | 7.064 |
| **95-th percentyl** | 33.9805 |
| **Maksimum** | 45199.914 |
| **Rozstęp** | 45199.643 |
| **Rozstęp miedzykwartylowy (IQR)** | 4.35 |

Statystyki opisowe

|  |  |
| --- | --- |
| **Odchylenie standardowe** | 323.54818 |
| **Współczynnik wariancji (CV)** | 12.763052 |
| **Kurtoza** | 2617.7329 |
| **Średnia** | 25.350377 |
| **Median Absolute Deviation (MAD)** | 1.327 |
| **Skośność** | 35.789048 |
| **Sum** | 27378686 |
| **Wariancja** | 104683.43 |

Wartości najczęstsze

| **Wartość** | **Liczebność** | **Częstotliwość (%)** |
| --- | --- | --- |
| 2.688 | 456 | <0.1% |
| 2.665 | 456 | <0.1% |
| 2.675 | 456 | <0.1% |
| 2.714 | 453 | <0.1% |
| 2.694 | 447 | <0.1% |
| 2.732 | 447 | <0.1% |
| 2.707 | 443 | <0.1% |
| 2.678 | 442 | <0.1% |
| Pozostałe wartości | 1075536 | 95.9% |
| (Braki danych) | 35784 | 3.7% |

Wartości ekstremalne - 10 wartości maksymalnych

|  |
| --- |
|  |
| **Wartość** | **Liczebność** | **Częstotliwość** (%) |
| 45199.914 | 1 | < 0.1% |
| 45199.142 | 1 | < 0.1% |
| 45198.53 | 1 | < 0.1% |
| 43611.586 | 1 | < 0.1% |
| 43593.418 | 4 | < 0.1% |
| 41089.046 | 1 | < 0.1% |
| 39626.126 | 3 | < 0.1% |
| 16467.662 | 1 | < 0.1% |
| 16464.58 | 1 | < 0.1% |
| 15144.021 | 1 | < 0.1% |

## **Zmienna - StartDateTimeUtc** - charakterystyki zmiennej

|  |  |
| --- | --- |
| Wartości unikatowe | 1107767 |
| Wartości unikatowe **(%)** | 98.8% |
| Braki danych | 0 |
| Braki danych **(%)** | 0.0% |
| **Minimum** | 2023-02-01 00:06:11.107000 |
| **Maksimum** | 2023-03-31 23:53:31.460000 |

Obraz zawierający tekst, Czcionka, zrzut ekranu, diagram

Opis wygenerowany automatycznie

## **Tablice korelacji** – korelacja między zmiennymi

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **InputPageCount** | **MessagesInQueue** | **Retries** | **ProcessingDuration\_sec** | **Duration\_sec** | **Queue** |
| **InputPageCount** | 1.000 | 0.002 | 0.015 | 0.265 | 0.222 | 0.000 |
| **MessagesInQueue** | 0.002 | 1.000 | 0.071 | 0.411 | 0.472 | 0.037 |
| **Retries** | 0.015 | 0.071 | 1.000 | 0.059 | 0.107 | 0.035 |
| **ProcessingDuration\_sec** | 0.265 | 0.411 | 0.059 | 1.000 | 0.889 | 0.003 |
| **Duration\_sec** | 0.222 | 0.472 | 0.107 | 0.889 | 1.000 | 0.032 |
| **Queue** | 0.000 | 0.037 | 0.035 | 0.003 | 0.032 | 1.000 |

Obraz zawierający tekst, zrzut ekranu, diagram, kwadrat

Opis wygenerowany automatycznie

## **Tablice korelacji** – korelacja między brakami danych

Obraz zawierający tekst, zrzut ekranu, diagram, linia

Opis wygenerowany automatycznie