1. **Instrukcja uruchomienia aplikacji**

Aby uruchomić aplikację wystarczy na komputerze z systemem Windows i dostępem do sieci Internet uruchomić załączony plik instalacyjny „setup.exe”. Po zainstalowaniu aplikacji należy uruchomić plik wykonywalny „ELT.exe”.

1. **Opis funkcjonalności**

**Funkcjonalności procesu ETL:**

1. Proces pozyskania danych (E – Extract) - najpierw nacisnąć przycisk „EXTRACT”, co uruchomi procedurę „web crawlingu” strony [www.ocen-piwo.pl](http://www.ocen-piwo.pl), gdzie zbierane są najpierw dane o najwyżej ocenianych piwach przez użytkowników strony. Proces realizowany jest poprzez parsowanie informacji o piwach z odpowiedniej zakładk, gdzie na każdej stronie parsowanych jest 10 piw. Proces powtarza się dla pierwszych 10 stron, co pozwala zatem zebrać dane na temat 100 najlepiej ocenianych na rynku piw w Polsce. Kolejnym uruchomionym procesem jest parsowanie komentarzy dla wszystkich zebranych przed chwilą piw. W pierwszym etapie program zebrał niezbędne linki url do wszystkich podstron. Na etapie parsowania opinii użytkowników program otwiera podstronę każdego piwa i zbiera dane takie jak: nazwa użytkownika, wystawiona przez niego ocena, opinia oraz czas dodania wpisu. Ponieważ w drugim etapie otwierane jest 100 podstron proces może w zależności od szybkości łącza internetowego oraz aktualnego obciążenia strony ocen-piwo.pl potrwać od kliku do kilkunastu minut.
2. Proces przekształcenia danych (T – Transform) – po pomyślnym wykonaniu procesu pozyskania danych odblokowany zostaje etap transformacji. Aby go rozpocząć użytkownik musi kliknąć przycisk „Transform”. Na etapie tym program przygotowuje dane do prawidłowego zapisu. Dane otrzymują jednoznacznie identyfikujące je identyfikatory, tak aby nie zostały niepotrzebnie zduplikowane.
3. Proces ładowania danych (L – Load) – po wykonaniu procesu z podpunktu b. można przejść do ostatniego kroku procesu ETL. Dane zostają zapisane do lokalnej bazy danych.

**Cały proces ETL można również wykonać za pośrednictwem jednego przycisku „ETL”.** Po jego kliknięciu wszystkie 3 kroki zostaną wykonane jeden po drugim bez konieczności dalszej interakcji użytkownika.

**Zapis do pliku csv:**

Po poprawnym wykonaniu wszystkich kroków z procesu ETL użytkownik ma możliwość zapisu zebranych danych w przystępnej postaci arkusza kalkulacyjnego csv. Aby rozpocząć proces zapisu danych wystarczy nacisnąć przycisk „SAVE TO .CSV FILE”. Program rozpocznie procedurę zapisu z podziałem na 3 pliki:

- Beers.csv (gdzie zapisane zostaną dane o każdym piwie)

- UserComments.csv (opinie zalogowanych użytkowników)

- GuestComments.csv (opinie niezalogowanych użytkowników – gości strony)

Podział na opinie zalogowanych i niezalogowanych użytkowników został wymuszony przez politykę strony ocen-piwo.pl, która dla użytkownika gościa nie pozwala na zostawienie oceny w postaci wartości liczbowej, a jedynie na słowna recenzję trunku.

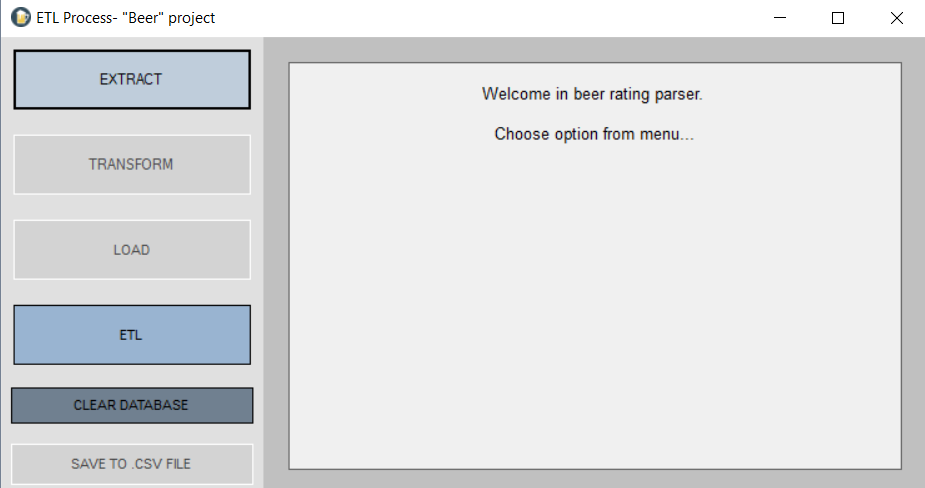
1. **Opis scenariuszy użycia aplikacji**

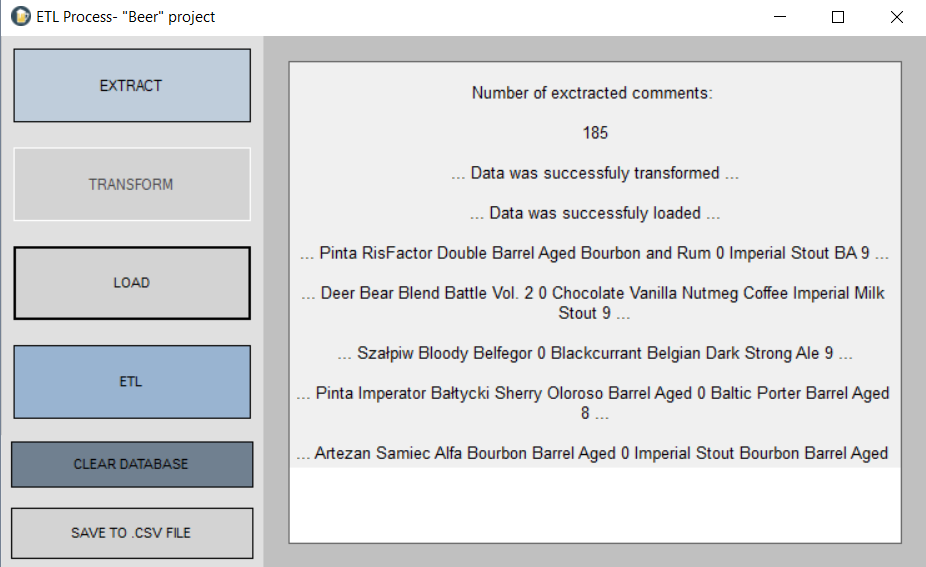
Aplikacja mogłaby przydać się firmom produkujących piwo w Polsce do analizy rynku. Dzięki naszej aplikacji analitycy tych firm mogliby w łatwy sposób zebrać opinie o najlepiej ocenianych przez użytkowników piwach. Dowiedzieć się jakie jakiego rodzaju piwa (z jakim procentowym udziałem alkoholu, mocą ekstraktu są aktualnie lubiane przez użytkowników). Posiadając takie informacje producenci mogliby np. rozpocząć prace nad wprowadzeniem na rynek piw o podobnej charakterystyce do tych najlepiej ocenianych, a także odnieść się do przedstawionych przez użytkowników opinii by urozmaicić lub jeszcze ulepszyć piwa o już dobrej opinii wśród użytkowników.

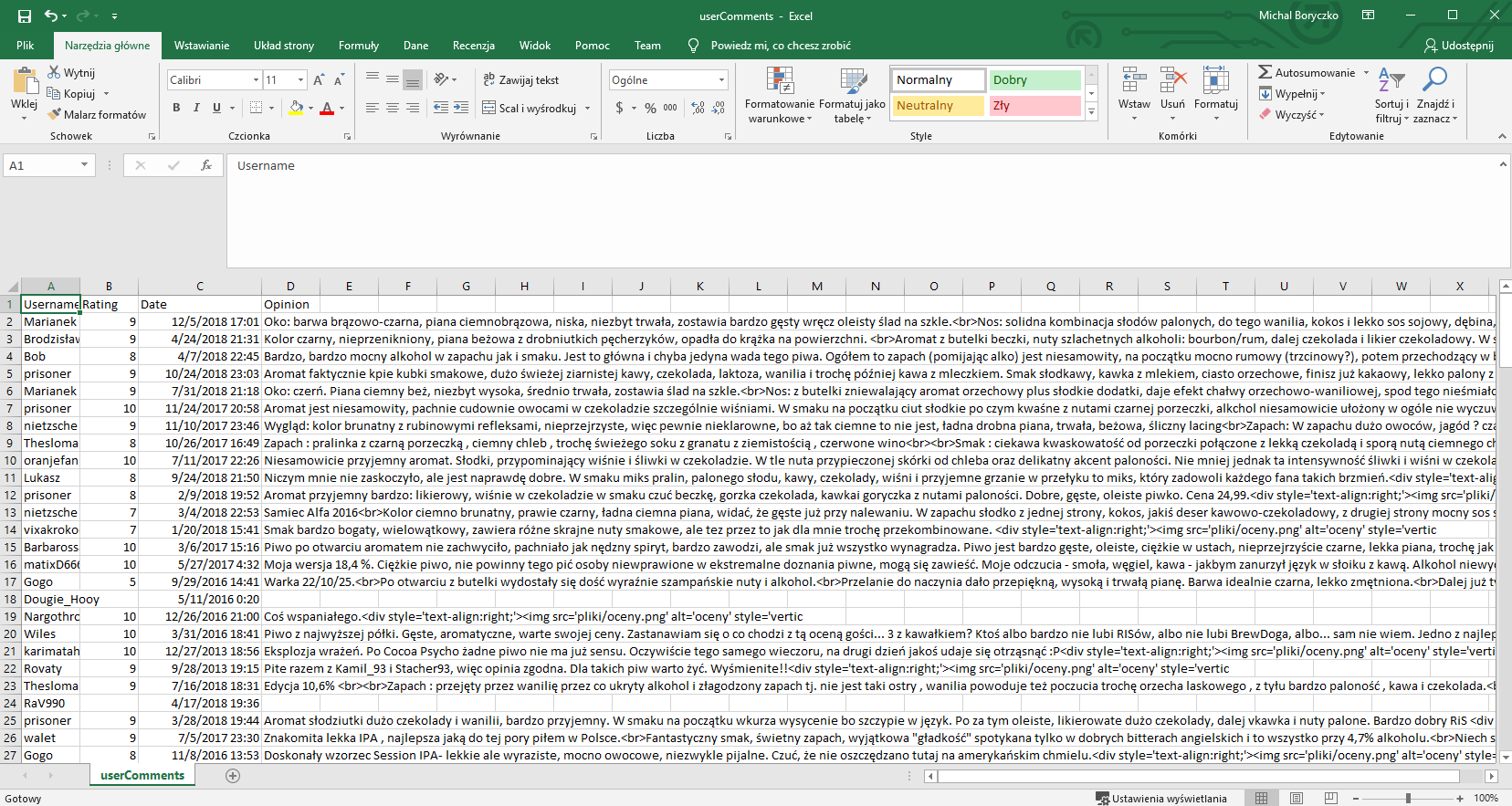
1. **Opis menu i widoków okna aplikacji**

Aplikacja ma bardzo prosty i intuicyjny interfejs dla użytkownika. Składa się z 6 przycisków:

* Extract – realizująca funkcjonalność pozyskania danych
* Transform – transformacja danych
* Load – załadowanie danych do lokalnej bazy danych
* ETL – proces realizujący wszystkie 3 kroki na raz
* CLEAR DATABASE – wyczyszczenie całej bazy danych
* SAVE TO .CSV FILE – zapis danych do arkusza kalkulacyjnego .csv





Rezultat zapisanych opinii użytkowników do pliku csv: