

Julia Błaszczyk

Stażysta inżynierii danych

✉ julia.blaszczyk@example.pl ☎ +48 500 209 768 📍 Toruń

DOŚWIADCZENIE ZAWODOWE

Stażysta inżynierii danych

07/2024 - 06/2025

Wisła Logistyka Miejska

Toruń

- Zbudowała w Pythonie import 18 plików GTFS dla 640 tys. dziennych przejazdów; po przeglądzie z 2 mentorami raport punktualności powstawał w 6 zamiast 40 minut.
- Dodała testy Great Expectations dla przystanków i godzin odjazdu; 97,8% rekordów przechodziło kontrolę w 5 kolejnych tygodniowych zrzutach.
- Przygotowała model DuckDB dla analityczki miejskiej obejmujący 14 linii autobusowych; ręczne łączenie tabel zajmowało 12 minut zamiast 90.
- Opisała błędy geokodowania na próbce 32 tys. przystanków i zgłosiła je dostawcy; liczba nieprzypisanych lokalizacji spadła o 21% w następnym imporcie.

Asystentka projektu danych

09/2023 - 06/2024

Toruńskie Laboratorium Danych

Toruń

- Napisała skrypt walidujący 11 tys. zgłoszeń laboratoryjnych podczas zajęć zespołowych; prowadzący odrzucał 8 typów błędnych dat przed analizą.
- Znormalizowała nazwy 4 200 ulic w projekcie uczelnianym z trzema studentami; udział dopasowanych adresów wzrósł z 88 do 96%.
- Dodała README i przykładowe zapytania SQL do repozytorium 6 osób; każdy członek odtwarzał wynik na tym samym zbiorze w mniej niż 15 minut.
- Zautomatyzowała pobieranie danych pogodowych z 5 stacji na potrzeby pracy dyplomowej; ręczne przygotowanie wykresów skróciła z 2 godzin do 18 minut.

PROJEKTY

Otwarte dane komunikacji miejskiej

Produkt danych dla obszaru: rozkłady jazdy i punktualność transportu. Obejmuje kontrolowane źródła, wersjonowane transformacje i sprawdzalne reguły jakości.

Technologie: Python, DuckDB, Airflow, Great Expectations

UMIEJĘTNOŚCI

Platformy i języki: Python, DuckDB, Airflow, Great Expectations, Git

Jakość i utrzymanie: Modelowanie danych, Testy, CI/CD, Monitoring, Dokumentacja

Współpraca: Ustalanie wymagań, Przegląd kodu, Obsługa incydentów, Dostarczanie między zespołami

WYKSZTAŁCENIE

Licencjat, informatyka w trakcie studiów rozkłady jazdy i punktualność transportu

10/2022 - 06/2026

Uniwersytet Mikołaja Kopernika w Toruniu

Toruń

CERTYFIKATY

Databricks Lakehouse Fundamentals

03/2025

Organizacja certyfikująca

Warsztat jakości danych: rozkłady jazdy i punktualność transportu

10/2023

Polskie Forum Inżynierii Danych

JĘZYKI

Polski: Język ojczysty

Angielski: Poziom zawodowy