TARTU ÜLIKOOL

Arvutiteaduse instituut

Infotehnoloogia mitteinformaatikutele õppekava

Merlin Raud, Henri Pokk, Kairi Käiro

**Eesti liiklusõnnetustest aastatel 2011-2022**

Grupi projekt

TARTU 2023

**Tutvustus**

Meie grupp tegi projekti Eesti liiklusõnnetustest aastatel 2011-2022. Andmestiku saime lehelt: <https://avaandmed.eesti.ee/datasets/inimkannatanutega-liiklusonnetuste-andmed>. Andmestikule on võimalik ligi pääseda ka APIga aga meie laadisime andmebaasi alla ja töötasime sellega Python notebook keskkonnas kasutades Visual Studio Code-i. Lisaks liiklusõnnetuste andmetele kasutasime ilmajaama andmeid ja liitsime kogu andmetikule juurde temperatuuri väärtused**. Meie projekti eesmärgiks** on anda ülevaade liiklusõnnetuste statistikast Eestis aastatel 2011-2022. Meie loodud töölaud (*dashboard*) on dünaamiline, mis tähendab, et kasutajal on võimalus valiku nuppude abil ehitada üles enda vajadustele vastav päring. Elementide paigutamisel töölauale jälgisime põhimõtet, et kõige olulisel informatsioon, nt liiklusõnnetuste, hukkunute ja vigastatute koguarv oleks esitatud kõige suuremalt ja haaraks lugeja tähelepanu. Esimene oluline joonis on Eesti kaart, kus on tähistatud liiklusõnnetuste asukohad. Järgnevad joonised liiklusõnnetuste statistikast (õnnetuste arv erinevatel ilmastikuoludel, õnnetuste arv erinevatel tee elementidel, erinevat liiki liiklusõnnetuste arv, õnnetuste arv aastate kaupa) esitasime joon- ja tulpdiagrammina koos valiku nuppudega.

**Andmete puhastamine**

Üldine andmete kvaliteet oli hea. Enamus lahtreid oli täidetud ja kui puudus info mõne elemendi kohta, siis see oli märgitud NaN-ina. Veerud olid numbrilised ja kategoorilised. Andmestikus oli üle 50 erineva veeru, millest valisime endale vajaminevad edasiseks andmetöötluseks välja. Kõiki veerge jooniste ja graafikute tegelemisel ei kasutanud, sest osad dubleerisid infot. **Koordinaadid** olid esitatud EPSG:3301 süsteemis, mis tuli ümber teisendada geograafilisteks (EPSG:4326). Mõningad liiklusõnnetuste asukohad olid väljaspool Eestis, näiteks Lätis, Tiibetis ja Aafrikas. Need eemaldasime edasisest andmeanalüüsist. Samuti tuli **piirkiiruse** veeru andmeid korrastada, sest seal esines mittetäisarvulisi väärtusi. Andmete puhastamise tegime Python notebookis. Edasine töö käis puhastatud andmetega.

**Töölaua tegemise protsess**

Töölaua jaoks vajaminevate valiku nuppude, jooniste tegemiseks kasutasime Plotly Phyton (plotly v 5.10.0 ) lahendust. Kogu töölaua liitsime kokku Python Panelis (panel v 0.14.2). Kogu koodi käivitamiseks on vaja installida järgevad paketid:

pandas==2.0.1

numpy==1.23.1

seaborn==0.12.1

geopandas==0.13.0

matplotlib==3.6.2

plotly==5.10.0

holoviews==1.15.3

numexpr==2.8.4

panel==0.14.2

bokeh==2.4.3

pyproj==3.4.0

shapely==2.0.1

folium==0.12.1.post1

Töötasime Studio Visual Code keskkonnas, mille kaudu laadisime kogu koodi ja sellega seonduvad failid GitHubi. Enne töölaua tegemise protsessi tegime Figmas UX disaini joonise, milline peaks lõplik töölaua versioon välja nägema.

*Töölaua esimene veerg*

Töölaual on kasutajal võimalik valida teda huvitav ajavahemik, mis on esitatud liugurina (*slider*). Valides sobiva ajavahemiku, joonistel olev informatsioon muutub automaatselt.

Iga joonise kohta on esitatud erinevad valikunupud, mis on paigutatud joonise üles. Liiklusõnnetuste osalised on esitatud töölaua esimese veeru üleval, mida on võimalik kustutada ja uuesti valida. Tehtud valik mõjutab terve töölaua erinevaid jooniseid.

Liiklusõnnetuste joonisel on võimalik valida, kas kasutaja soovib näha liiklusõnnetuse arvu aastate, kuude, päevade ja tundide kaupa. Automaatselt esitatakse töölaual esimene nupu valik ehk antud hetkel liiklusõnnetuste arv aastate lõikes.

*Töölaua teine veerg*

Töölaua teise veeru esimene joonis on Eesti kaart, millele on kantud kõik liiklusõnnetused aastatel 2011-2022. Kaarti on võimalik sisse-välja suumida nii hiire kui klaviatuuriga. Liuguriga on võimalik selekteerida huvipakkuv periood välja ja kaardil esitatakse vastava perioodi liiklusõnnetused.

Õnnetuste osalejate arvu/õnnetuse liigi joonisel on kasutajal võimalik valida, kas soovib joonisel näha õnnetuste osalejate arvu jaotust või õnnetuste liiki. Seda saab teha klikkides vastava valiku nupu peal.

*Töölaua kolmas veerg*

Töölaua kolmanda veeru esimene joonis on seotud mitmete valikutega. Esimene ja automaatselt rakenduv valik näitab õnnetuste arvu erinevatel tee elementidel. Kasutajal on võimalik valida järgnevate valikute vahel: õnnetuste arv erinevate tee objektidel, õnnetuste arv seotuna tee kurvilisusega, õnnetuste arv seotuna tee tasasusest, õnnetuste arv seotuna tee seisundist, õnnetuste arv seotuna teekattest ja õnnetuste arv seotuna piirkiirusega.

Töölaua viimasel joonisel on võimalik vaadata õnnetuste arvu mõjutatuna erinevate keskkonnaga seotud parameetrite kaudu: ilmastikuolud, valge aeg, valgustus, temperatuur ja teekatte seisund.

**Tekkinud probleemid**

Loodud töölaud ei läinud kõikidel grupiliikmetel arvutites käima ja esines erinevaid veateateid. Vahest aitas, kui vajaminevad paketid uuesti alla laadida. Oli ka olukord, kus kõik elemendid töölaual eraldi läksid käima aga terve töölaua kokkupaneku kood ei töötanud. Ilmselt oli probleemiks varasemate pakettide alla laadimisest tekkinud konflikt. Antud koodiga töötasime Visual Studio Code programmis ja üks grupiliige töötas Jupiter colab keskkonnas.