



TALLINNA TEHNIKAÜLIKOOL  
INSENERITEADUSKOND  
Virumaa kolledž

**TalTech Virumaa kolledži vastuvõtuandmete  
visualiseerimine Power BI tarkvara abil**

**Visualization of TalTech Virumaa College Admission Data Using  
Power BI Software**

ARUKAD SÜSTEEMID JA RAKENDUSINFOTEHNOLOOGIA ÕPPEKAVA  
LÕPUTÖÖ

Üliõpilane: Pauliina Kolde

Üliõpilaskood: 182732EDTR

Juhendaja: Olga Dunajeva,  
vanemlektor

Kaasjuhendaja: Valeria Juštšenko,  
haridustehnoloog

## **AUTORIDEKLARATSIOON**

Olen koostanud lõputöö iseseisvalt.

Lõputöö alusel ei ole varem kutse- või teaduskraadi või inseneriplomit taotletud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on viidatud.

# LIHTLITSENTS LÕPUTÖÖ ÜLDSUSELE KÄTTESAADAVAKS TEGEMISEKS JA REPRODUTSEERIMISEKS<sup>1</sup>

Mina Pauliina Kolde (sünnikuupäev: 21.02.1991)

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose TalTech Virumaa kolledži vastuvõtuandmete visualiseerimine Power BI tarkvara abil, mille juhendaja on Olga Dunajeva,
  - 1.1. reprodutseerimiseks säilitamise ja elektroonilise avaldamise eesmärgil, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
  - 1.2. üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. Olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. Kinnitan, et lihtlitsentsi andmisega ei rikuta kolmandate isikute intellektuaalomandi ega isikuandmete kaitse seadusest ja teistest õigusaktidest tulenevaid õigusi.

---

<sup>1</sup> Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingulise tegevusega ning lõputöö kaas- või ühisautori(d) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtajaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.

# SISUKORD

EESSÕNA .....	6
SISSEJUHATUS .....	7
1. PROBLEEMI KIRJELDUS JA TARKVARA VALIK .....	8
2. ANDMETE ETTEVALMISTAMINE .....	9
2.1 Esialgne andmestik.....	9
2.2 Andmestiku korrastamine MS Excelis .....	10
2.3 Andmestiku ettevalmistamine Power Query Editoriga .....	10
2.4 Andmemudeli loomine.....	11
2.4.1 Dimensioonitabeli DimEriala loomine .....	12
2.4.2 Dimensioonitabeli DimÜliõpilane loomine .....	13
2.4.3 Faktitabeli FactTabel loomine .....	13
2.4.4 Andmemudeli seoste loomine Power BI-s.....	14
2.5 DAX keeles arvutatud mõõdikute ja veergude lisamine.....	15
3. VISUALISEERIMINE .....	17
3.1 Kaardi visuaalide loomine .....	17
3.2 Ringdiagrammide loomine .....	18
3.3 Sektordiagrammide loomine .....	19
3.4 Tabeli visuaali loomine .....	20
3.5 Joon- ja virnastatud tulpdiagrammide loomine .....	21
3.6 Rühmitatud tulpdiagrammide loomine.....	24
3.7 100% virnastatud tulpdiagrammi loomine .....	25
3.8 Geograafilise kaardi visuaali loomine .....	25
3.9 Tornaadodiagrammi loomine.....	26
4. ARUANDE LOOMINE.....	28
4.1 Aruande kujundus .....	28
4.2 Aruande päise loomine.....	30
4.3 Aruande leheküljed.....	34
4.3.1 Üldine vastuvõtustatistika .....	34
4.3.2 Vastuvõtu trendid .....	35
4.3.3 Rakendusõppe vastuvõtt.....	36
4.3.4 Magistriõppe vastuvõtt .....	37
4.3.5 Üliõpilaste sooline jaotus .....	38

4.3.6	Üliõpilaste vanuseline jaotus .....	39
4.3.7	Üliõpilaste päritolu .....	40
4.3.8	Üliõpilaste eelnev õppeasutus.....	41
5.	RAKENDUSE LOOMINE JA AUTOMATISEERIMINE .....	43
5.1	Faili algallika muutmine OneDrive'i asukohaks .....	43
5.2	Aruande publitseerimine rakendusena .....	44
5.3	Rakenduse uuendamise automatiseerimine.....	46
5.4	Rakenduse uuendamise protsessi kirjeldus .....	48
	KOKKUVÕTE .....	49
	SUMMARY .....	51
	KASUTATUD KIRJANDUSE LOETELU .....	52

## EESSÕNA

Käesoleva lõputöö teema valik ja sõnastus on tehtud töö autori Pauliina Kolde ja juhendaja Olga Dunajeva koostöös, lähtudes töö autori soovist õppida kasutama Power BI visualiseerimise tarkvara ja TalTech Virumaa kolledži vajadusest igal õppeaastal visualiseerida ülikooli sisseastunute andmeid.

Lõputöös kasutatud andmestik pärineb TalTech Virumaa kolledži haridustehnoloog Valeria Juštšenkolt, kes kaardistas kolledži vajadused rakendusele ja kelle jagatud esitlusmaterjalid 2021. ja 2023. aasta TalTech Virumaa kolledži vastuvõtustatistikast olid sisendiks uue rakenduse väljatöötamisel.

Autor soovib tänada juhendajat, TalTech Virumaa kolledži lektorit Olga Dunajevat teema valiku, konsultatsioonide ja juhendamise eest, kaasjuhendajat haridustehnoloog Valeria Juštšenkot aruande loomiseks sisendi andmise eest ja enda perekonda ning lähedasi toetuse eest.

## SISSEJUHATUS

Infoühiskonna andmemahtude suurenedes on andmeanalüüs muutumas järjest olulisemaks oskuseks. Andmete analüüsimise lahutamatuks osaks on kujunenud andmete visualiseerimine, mis hõlbustab andmetest aru saamist. TalTech Virumaa kolledžis visualiseeritakse paremaks analüüsimiseks vastuvõtuandmeid igal õppeaastal, kasutades selleks MS Office tarkvara.

Käesoleva lõputöö peamiseks eesmärgiks oli TalTech Virumaa kolledži vastuvõtuandmete visualiseerimise automatiseerimine MS Power BI tarkvara abil, et hõlbustada kolledži iga-aastast vastuvõtuandmete visualiseerimise protsessi.

Teema valikul lähtuti töö autori huvist andmeteaduse vastu ja TalTech Virumaa kolledži järjepidevast vajadusest vastuvõtuandmete töötlemiseks. Power BI tarkvara valiti andmete töötlemise automatiseerimiseks, lähtudes tema turupositsioonist, kättesaadavusest Tallinna Tehnikaülikooli töötajatele ja TalTech Virumaa kolledži õppekavast.

Lõputöö ülesannete raames korrastati TalTech Virumaa kolledži aastate 2019–2023 vastuvõtu andmestik MS Excelis ja MS Power BI Desktopi tarkvaras, kasutades selleks Power Query Editori redaktorit. Seejärel visualiseeriti andmed ja koostati kaheksast leheküljest koosnev Power BI aruanne, mis publitseeriti rakendusena.

Käesoleva lõputöö raames antakse TalTech Virumaa kolledži vastuvõtuandmete Power BI tarkvara abil visualiseerimise protsessist põhjalik ülevaade. Esimeses peatükis kirjeldatakse probleemi ja tarkvara valikut lähemalt. Teine peatükk annab ülevaate andmete korrastamise, puhastamise ja Power Query Editoriga ette valmistamise protsessist kuni andmemudeli loomise ja visualiseerimiseks oluliste arvutatud mõõdikute ning veergude lisamiseni. Kolmandas peatükis kirjeldatakse aruandes kasutatud visuaalide loomist, neljandas peatükis kirjeldatakse aruande loomise protsessi ja antakse ülevaade kõigist aruande lehekülgedest ning neile loodud visuaalidest. Viies peatükk kirjeldab aruande publitseerimist rakendusena, mis uueneb automaatselt ja annab ülevaate rakenduse uuendamise protsessist.

Võtmesõnad: Power BI tarkvara, Power Query Editor, andmete visualiseerimine, aruande loomine, bakalaureusetöö.

# 1. PROBLEEMI KIRJELDUS JA TARKVARA VALIK

Andmestunud infoühiskonda iseloomustab inimeste igapäevase elu andmetest läbi põimumine. Digitehnoloogiate kasutuselevõtt on toonud kaasa ühiskonna andmestumisprotsessi, kus inimeste igapäevaseid tegevusi ning ühiskondlikke protsesse on võimalik teisendada arvulisteks, kvantitatiivseteks, süstematiseeritud ja analüüsitavateks andmeteks. [1]

Ajastul, kus andmed kogunevad justkui ise, kuid kasvab vajadus neid kirjeldada, kontrollida, analüüsida ja esitada lihtsasti loetaval kujul, on andmeteandus kiiresti arenev valdkond [2]. Kuna andmed omavad ärimaailma otsustusprotsessides kriitilist tähtsust, arendatakse järjepidevalt tööriistu ja strateegiaid andmete kogumiseks, analüüsimiseks ja visualiseerimiseks. Vastavalt prognoosidele on Business Intelligence'i (BI) visuaalsete tööriistade ja analüütika kasutuselevõtt järgnevatel aastatel kasvutrendis. [3]

TalTech Virumaa kolledž võtab igal aastal vastu uusi üliõpilasi ja, selleks et kolledži juhtimise strateegiaid välja töötada, on kolledžil tarvis omada täpset ülevaadet vastuvõtustatistika kohta, analüüsida vastuvõtu trende ja mõista üliõpilaste demograafilist tausta. See tähendab, et vastuvõtuandmeid tuleb igal aastal paremaks analüüsimiseks visualiseerida. Seni on vastuvõtuandmete visualiseerimine olnud TalTech Virumaa kolledžis töötavate haridustehnoloogide tööülesandeks, kes on kasutanud andmete visualiseerimiseks MS Exceli, MS Wordi ja MS PowerPointi tarkvara.

TalTech Virumaa kolledži vastuvõtuandmete visualiseerimine BI tarkvara abil võimaldab selle iga-aastase andmete visualiseerimise protsessi automatiseerida. BI tarkvara lahenduste pakkujaid on turul palju, kuid turuliidriteks on 2023. aastal uuringu- ja konsultatsioonifirma Gartner aruande järgi Microsoft Power BI ja Tableau tarkvaralahendused [4], mis on mõlemad ühtlasi ka TalTech Virumaa kolledžis õpetatavad BI tarkvarad.

Tableau tarkvaral on küll Power BI tarkvaraga võrreldes suurte andmekogumite puhul parem käsitlemine ja laiem platvormideülene tugi, kuid kuna ta on üldiselt kallim kui Power BI, siis väikeettevõtted kipuvad eelistama MS Power BI tarkvara, mille integreerimine teiste Microsofti toodetega on sujuvam [5].

Power BI tarkvara valiti andmete töötlemise automatiseerimiseks lähtudes sellest, et Tallinna Tehnikaülikoolis on samuti andmete visualiseerimiseks kasutusele võetud just Power BI tarkvaralahendus. Kuna Power BI aruannetele saab vajadusel anda juurdepääsu kõigile Tallinna Tehnikaülikooli töötajatele, siis sai see tarkvara valikul määravaks aspektiks. Lisaks kinnitas valikut see, et Power BI on Forbes.com andmetel 2023. aastal andmete visualiseerimiseks parim Business Intelligence'i tarkvara [6].



## 2. ANDMETE ETTEVALMISTAMINE

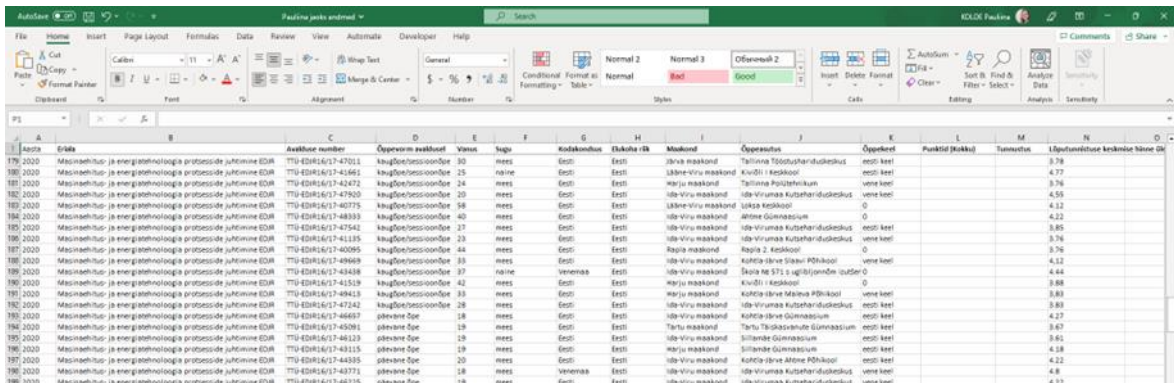
Andmete ettevalmistamine on andmeanalüüsi protsessis kriitilise tähtsusega etapp. Andmete kvaliteet määrab otseselt raportite ja analüüside usaldusväärsuse. Korrastamata andmed võivad põhjustada vigu, eksitavaid tõlgendusi ja ebatäpset analüüsi, samuti aeglustada raporteid ja vähendada nende reageerimiskiirust. Seetõttu on väga oluline tagada, et andmed oleksid enne analüüsi Power BI platvormil põhjalikult korrastatud. [7]

Käesoleva lõputöö raames algas TalTech Virumaa kolledži vastuvõtuandmete ettevalmistamise protsess esialgse andmestikuga tutvumisest. Seejärel tehti andmestikus mõningad parandused MS Excelis, mille järel toimus andmestiku korrastamine ja andmemudeli loomine Power BI Desktopi Power Query Editori redaktoris. Kui andmed olid korrastatud, siis loodi nende põhjal Power BI Desktopis andmete visualiseerimiseks olulised arvutatud mõõdikud ja veerud.

### 2.1 Esialgne andmestik

Käesolev lõputöö põhineb TalTech Virumaa kolledži aastate 2019–2023 vastuvõtuandmetel, mis on pärit Haridus- ja Teadusministeeriumi sisseastumise infosüsteemist (SAIS). Kuna SAISis olevad andmed sisaldavad üliõpilaste isikuandmeid, lepitati lõputöö raames kokku, et kasutatakse haridustehnoloogi loodud uut, eraldi Exceli faili isikustamata admetega, mida ei saa kindla üliõpilase isikuga siduda.

Haridustehnoloogi esitatud Exceli fail esialgse andmestikuga (vt Joonis 2.1) hõlmas 541 rida andmeid, mis olid jaotatud 15 veergu: Aasta, Eriala, Avalduse number, Õppevorm avaldusel, Vanus, Sugu, Kodakondsus, Elukoha riik, Maakond, Õppeasutus, Õppekeel, Punktid (kokku), Tunnustus, Lõputunnistuse keskmine hinne üle kõigi keskkooli ainete. Lisaks sisaldus üks nimetuseta veerg, mis sisaldas matemaatika eksami tulemuste väärtusi.



A	B	C	D	E	F	G	H	I	J	K	L	M	N
Aasta	Eriala	Avalduse number	Õppevorm avaldusel	Vanus	Sugu	Kodakondsus	Elukoha riik	Maakond	Õppeasutus	Õppekeel	Punktid (kokku)	Tunnustus	Lõputunnistuse keskmine hinne üle
179	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-47011	kuigapäevastõund	30	mees	Esti	Järva maakond	Tallinna Tööstushariduskeskus	vee keel	3,78		
180	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-41661	kuigapäevastõund	25	naine	Esti	Lääne-Viru maakond	Kivimäe keskkool	vee keel	4,77		
181	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-41672	kuigapäevastõund	24	mees	Esti	Järva maakond	Tallinna Põlvetehnikum	vee keel	3,76		
182	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-47920	kuigapäevastõund	20	mees	Esti	Ida-Viru maakond	Ida-Virumaa Kutsehariduskeskus	vee keel	4,55		
183	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-40775	kuigapäevastõund	58	mees	Esti	Lääne-Viru maakond	Lääne keskkool	0	4,12		
184	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-48859	kuigapäevastõund	40	mees	Esti	Ida-Viru maakond	Arme Gümnaasium	0	4,22		
185	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-47542	kuigapäevastõund	27	mees	Esti	Ida-Viru maakond	Ida-Virumaa Kutsehariduskeskus	vee keel	3,85		
186	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-41135	kuigapäevastõund	23	mees	Esti	Ida-Viru maakond	Ida-Virumaa Kutsehariduskeskus	vee keel	3,76		
187	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-40095	kuigapäevastõund	44	mees	Esti	Rajala maakond	Rajala 2. keskkool	0	3,76		
188	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-49669	kuigapäevastõund	33	mees	Esti	Ida-Viru maakond	Kõrge-äärne Põhikool	vee keel	4,12		
189	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-43438	kuigapäevastõund	37	naine	Venemaal	Ida-Viru maakond	Saia 16 STI s (agiliteeritud) ouster	0	4,44		
190	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-41519	kuigapäevastõund	42	mees	Esti	Järva maakond	Kivimäe keskkool	0	3,88		
191	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-48413	kuigapäevastõund	33	mees	Esti	Järva maakond	Kõrge-äärne Mäe Põhikool	vee keel	3,83		
192	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-47242	kuigapäevastõund	28	mees	Esti	Ida-Viru maakond	Ida-Virumaa Kutsehariduskeskus	vee keel	3,83		
193	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-46667	gümnaasium	18	mees	Esti	Ida-Viru maakond	Kõrge-äärne Gümnaasium	vee keel	4,27		
194	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-45061	gümnaasium	19	mees	Esti	Taru maakond	Taru Tööstus- ja Gümnaasium	vee keel	3,67		
195	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-48123	gümnaasium	19	mees	Esti	Ida-Viru maakond	Sillamäe Gümnaasium	vee keel	3,61		
196	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-43115	gümnaasium	19	mees	Esti	Järva maakond	Sillamäe Gümnaasium	vee keel	4,18		
197	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-44455	gümnaasium	20	mees	Esti	Ida-Viru maakond	Kõrge-äärne Põhikool	vee keel	4,22		
198	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-43771	gümnaasium	18	mees	Venemaal	Ida-Viru maakond	Ida-Virumaa Kutsehariduskeskus	vee keel	4,8		
199	2020	Masinehüüsi ja energiatehnoloogia protsesside juhtimine EDAR	TTU-EDAR16/17-48225	gümnaasium	19	mees	Esti	Ida-Viru maakond	Ida-Virumaa Kutsehariduskeskus	vee keel	4,22		

Joonis 2.1 Esialgne andmestik MS Excelis

Andmestikuga põhjalikumalt tutvudes ilmnes, et selles esineb mitmeid vigu. Esimesed parandused tehti Exeli failis ja andmete ettevalmistamise protsessi automatiseerimiseks kasutati Power BI Desktopi Power Query Editorit.

## 2.2 Andmestiku korrastamine MS Excelis

Sisseastumise statistika uuendatakse TalTech Virumaa kolledžis igal sisseastumise perioodil. Selleks et poleks probleemi Power BI aruannete uuendamise, lepitati haridustehnoloogiga kokku, et luuakse korrastatud Exceli fail, mis on Power BI aruande loomise ja uuendamise aluseks.

Microsoft Excelis viidi läbi korrastustöö, kus Eriala ja Õppevormi veergude nimetused viidi standardsele kujule. Ühe olulisema muudatusena muudeti Avalduse numbriga veerg Eriala koodiks, kasutades Exceli *Text to Columns* funktsiooni, et eraldada avalduse numbrist eraldi eriala kood. Puuduvad erialakoodid lisati manuaalselt vastavalt eriala nimetustele, ning see protsess toimis haridustehnoloogiga, kuna avalduse number oli kättesaadav vaid vanematel andmetel ja selle numbriga säilitamiseks andmestikus puudus vajadus.

Lisaks kustutati üksikud nullväärtused tekstiveergudest ning nimetuseta veerule anti nimi "Matemaatika eksam". Kõik muudatused on visuaalselt märgitud kollase värviga (vt Joonis 2.2).

Edaspidi Power BI aruandele andmete lisamiseks tuleb lisanduvad TalTech Virumaa kolledži vastuvõtuandmed lisada sellesse Exceli faili.

1	Aasta	Eriala kood	Eriala	Õppevorm	Vastu	Sugu	Kodakondsus	Elukoht	Maakond	Õppesutus	Õppekeel	Punktid (Rakus)	Tunnustus	Lõputunnistuse keskmine hinne üle kõigi keskkooli ainetest	Matemaatika eksam
179	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	30	mees	Eesti	Järva maakond	Tallinna Tööstushariduskeskus	estni keel		3,76			
180	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	25	naine	Eesti	Ida-Viru maakond	Kuusaloo	estni keel		4,77			
181	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	24	mees	Eesti	Haju maakond	Tallinna Polütehnikum	vene keel		3,76			
182	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	20	mees	Eesti	Ida-Viru maakond	Ida-Virumaa Kutsehariduskeskus	vene keel		4,55			
183	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	58	mees	Eesti	Järva-Viru maakond	Laiuse Keskkool	estni keel		4,12			
184	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	40	mees	Eesti	Ida-Viru maakond	Järve Gümnaasium	estni keel		4,22			
185	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	27	mees	Eesti	Ida-Viru maakond	Ida-Virumaa Kutsehariduskeskus	estni keel		3,85			
186	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	23	mees	Eesti	Ida-Viru maakond	Ida-Virumaa Kutsehariduskeskus	vene keel		3,76			
187	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	44	mees	Eesti	Rapla maakond	Rapla 2. Keskkool	estni keel		3,76			
188	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	33	mees	Eesti	Ida-Viru maakond	Kõrva-Järve Steapi Põhikool	vene keel		4,12			
189	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	37	naine	Venemaa	Ida-Viru maakond	Skõta Nr 371 s ühisgümnaasium (tutustajem angli ja idia)	estni keel		4,44			
190	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	42	mees	Eesti	Haju maakond	Kuusaloo	estni keel		3,88			
191	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	33	mees	Eesti	Haju maakond	Kõrva-Järve Mõnva Põhikool	vene keel		3,83			
192	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	sessioonõpe	28	mees	Eesti	Ida-Viru maakond	Ida-Virumaa Kutsehariduskeskus	estni keel		3,83			
193	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	päevaõpe	18	mees	Eesti	Ida-Viru maakond	Kõrva-Järve Gümnaasium	estni keel		4,27			
194	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	päevaõpe	19	mees	Eesti	Tartu maakond	Tartu Täiskasvanute Gümnaasium	estni keel		3,67			
195	2020	EDIR	Masinaehitus ja energiatehnoloogia protsesside juhtimine	päevaõpe	19	mees	Eesti	Ida-Viru maakond	Sillamäe Gümnaasium	estni keel		3,61			

Joonis 2.2 Parandustega andmestik MS Excelis

## 2.3 Andmestiku ettevalmistamine Power Query Editoriga

Power Query Editor on Microsofti ärianalüütika tööriist, mille abil on võimalik andmeid puhastada, muuta ja ümber vormindada vastavalt vajadustele [8]. See tööriist on integreeritud mitmete Microsofti rakendustega, sh Power BI ja Exceliga, võimaldades

kasutajatel luua korduskasutatavaid andmeühenduvusi ja ettevalmistusloogikaid [9]. Andmestiku struktuuri korrastamise käigus Power Query Editoris automatiseeritakse korrastamisprotsess, tagades, et algfaali uute andmeridade lisamisel viiakse Power BI Desktopis läbi samad korrastamise toimingud.

Andmete korrastamise protsessi Power Query Editoris, tuues välja probleemid andmetega ja lahendused, mis neile leiti, kirjeldab Tabel 2.1:

Tabel 2.1 Andmete korrastamine Power Query Editoris

<b>Probleem</b>	<b>Lahendus</b>
Lõputunnistuse keskmise hinde veeru väärtustes olid kümnendkohad eraldatud nii punktide kui ka komadega.	Veeru väärtustes asendati komad punktidega.
Lõputunnistuse keskmise hinde veeru nimetus oli liiga pikk.	Veeru nimetus muudeti lühemaks.
Andmestikus puudus veerg, mis sisaldaks ainult unikaalseid väärtusi.	Lisati indekseeritud veerg.
Tabelisse olid importimise käigus tekkinud tühjad read ja veerud.	Tabelist eemaldati tühjad read ja veerud.
Teksti ja numbriveergudes esines määratlemata väärtuseid ehk nullväärtuseid.	Nullväärtused asendati teksti veergudes vaikeväärtusega "määramata" ja numbriveergudes nulliga.
Andmestikus esines erineval sõnakujul dubleeritud tekstiväärtuseid.	Andmestikus esinenud erineval sõnakujul dubleeritud tekstiväärtused korrastati, kasutades väärtuste asendamise funktsiooni.
Maakonna veerus osad maakonna nimetuse väärtused sisaldasid liidet „maakond“ ja teised mitte.	Maakonna veerg korrastati, poolitades veeru tühiku kohalt. Loodi uus kohandatud veerg väärtusega "maakond", mille järel liideti kaks veergu kokku.
Power Query Editori rakendatud etappide jaotis sisaldas automaatselt tekkinud ebavajalikke toiminguid ja vajab selgemat struktuuri.	Power Query Editori rakendatud etappide jaotisest eemaldati automaatselt tekkinud ebavajalikud toimingud ja selgema arusaamise huvides muudeti etappide nimetusi.

Andmestiku korrastamise käigus tehti mitmeid olulisi toiminguid, mille eesmärk oli tagada andmete kvaliteet ja ühtsus. Andmestiku ettevalmistamise tulemusena loodi 16 veerust ja 540 andmerekast koosnev tabel, mis oli aluseks andmemudeli koostamisel.

## 2.4 Andmemudeli loomine

Kuigi Power BI võimaldab töötada ka ühe suure andmetabeliga, siis soovitud praktika on viia andmed üle täheskeemi järgivale andmemudelile [10]. Täheskeemi põhjal andmemudeli ülesehitamise põhilisteks eelisteks on arvutuskiiruse optimeerimine ja kasutajamugavuse tagamine [11].

Täheskeemi põhimõtte seisneb selles, et andmed on jagatud dimensioonitabeliteks, mis

ühenduvad faktitabeliga. Faktitabel sisaldab dimensioonivõtmete veerge ja numbriliste väärtustega veerge. Kui faktitabel toetab summaarseid arvutuskäike, siis dimensioonitabelid sisaldavad kirjeldavaid andmeid, toetades filtreerimist ja rühmitamist. [12]

Käesoleva lõputöö andmemudeli struktuur koosneb kahest dimensioonitabelist ja ühest faktitabelist. Andmemudeli loomise protsessi on detailsemalt kirjeldatud järgnevas alapeatükkides.

### 2.4.1 Dimensioonitabeli DimEriala loomine

Andmemudeli koostamise esimene etapp hõlmas Erialaid dimensioonivõtme loomist, ühendades aasta, erialakoodi ja õppevormi esimese tähe väärtused uude veergu nimega Erialaid. Dimensioonivõti Erialaid võimaldas loodud DimEriala tabelist eemaldada dubleeritud andmerealad.

DimEriala tabelile lisati uus kohandatud veerg nimega Õppeaste, kuhu määrati väärtused sõltuvalt Erialaid koodi viimasest tähemärgist. Veerg Õppeaste lisati kasutades valemit: `if Text.End([Erialaid kood], 1) = "R" then "Rakenduskõrgharidus" else "Magister"`.

DimErialaid tabeli 6 veerust koosneva andmestiku struktuuri näitab Tabel 2.2:

Tabel 2.2 Dimensioonitabeli DimErialaid struktuur

Veerg	Andmetüüp	Kirjeldus
Erialaid (Võti)	Tekst	Dimensioonivõti, mis koosneb sisseastumise aastast, erialakoodist ja vastavalt õppevormile „s“ või „p“ tähisest.
Aasta	Täisarv	Aasta, mil sisseastumisavaldus esitati.
Erialaid kood	Tekst	Neljakohtaline kood erialaid identifitseerimiseks.
Erialaid	Tekst	Erialaid täispikk nimetus.
Õppevorm	Tekst	Erialaid õppevorm avaldusel („päevaõpe“ või „sessioonõpe“).
Õppeaste	Tekst	Erialaid õppeaste („Rakenduskõrgharidus“ või „Magister“).

Dimensioonitabelist duplikaatide eemaldamine võimaldas andmete hulka märgatavalt vähendada, nii et varasema 540 andmerealaga võrreldes koosneb DimErialaid tabel vaid 33 andmerealast.

Lisaks tagati Power Query Editori töö optimeeritus, kustutades tehtavate toimingute loetelust algandematega tehtavad toimingud, mis puudutasid veerge, mis said loodud dimensioonitabelist eemaldatud, ja jäeti alles vaid dimensioonitabeli DimErialaid loomiseks olulised toimingud.

## 2.4.2 Dimensioonitabeli DimÜliõpilane loomine

Dimensioonitabelisse DimÜliõpilane koondati kõik üliõpilasi kirjeldavad andmed ja loodi indeks-veerg, millest sai dimensioonivõti. Power Query Editori töö optimeerimiseks jäeti alles vaid toimingud, mis olid dimensioonitabeli DimÜliõpilane loomiseks olulised. Lõpptulemusena koosneb dimensioonitabel DimÜliõpilane 540 andmerekast ja 9 veerust, mille struktuuri näitab Tabel 2.3:

Tabel 2.3 Dimensioonitabeli DimÜliõpilane struktuur

Veerg	Andmetüüp	Kirjeldus
Index (Võti)	Täisarv	Unikaalne indeks igale andmerekale, vahemikus 1–540.
Vanus	Täisarv	Üliõpilase vanus aastates.
Sugu	Tekst	Üliõpilase sugu.
Kodakondsus	Tekst	Üliõpilase kodakondsuse teave.
Elukoha riik	Tekst	Riik, kus üliõpilane elab.
Maakond	Tekst	Maakond, kust üliõpilane pärit on.
Õppeasutus	Tekst	Viimati lõpetatud õppeasutuse nimi.
Õppekeel	Tekst	Viimati lõpetatud õppeasutuse õppekeel.
Tunnustus	Tekst	Kas üliõpilane lõpetas viimase õppeasutuse medaliga.

## 2.4.3 Faktitabeli FactTabel loomine

Fakti tabelisse FactTabel koondati dimensioonivõtmed ja arvutatavad väärtused, ning Power Query Editori optimeerimiseks kustutati toimingute loetelust algandmetega tehtavad toimingud, mis puudutasid veerge, mis faktitabelist kustutati. Faktitabelis FactTabel säilitati kõik 540 andmerida ja lõpptulemusena koosneb see viiest veerust, mille struktuuri näitab Tabel 2.4:

Tabel 2.4 Faktitabeli FactTabel struktuur

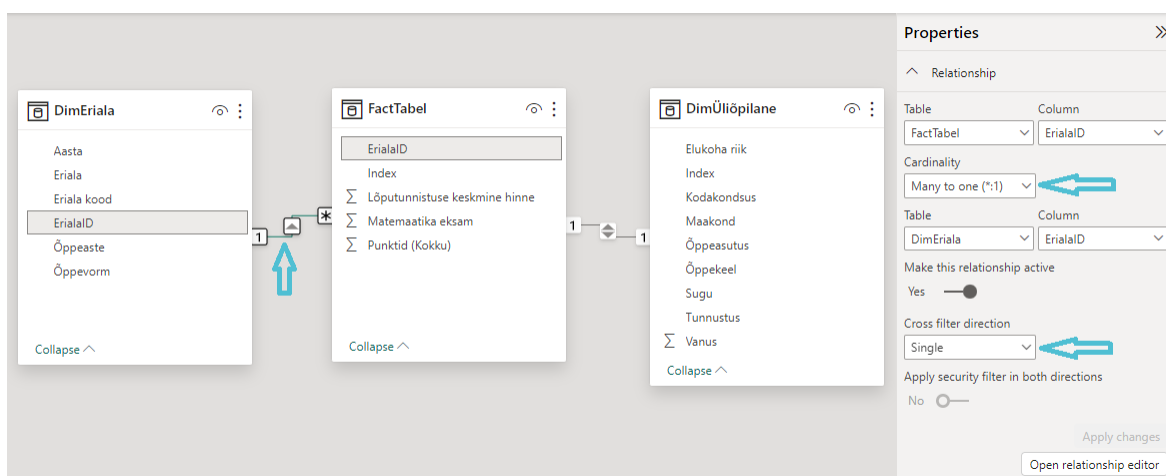
Veerg	Andmetüüp	Kirjeldus
Index (Võti)	Täisarv	Unikaalne indeks igale andmerekale, vahemikus 1–540.
ErialaID (Võti)	Tekst	Dimensioonivõti, mis koosneb sisseastumise aastast, eriala koodist ja vastavalt õppevormile tähisest „s” või „p”.
Punktid (Kokku)	Kümnendarv	Kokkuarvestatud punktide skoor TalTech Virumaa kolledžisse sissastumiseks.
Lõputunnistuse keskmine hinne	Kümnendarv	Viimati lõpetatud õppeasutuse keskmine hinne.
Matemaatika eksam	Kümnendarv	Matemaatika eksami tulemused.

## 2.4.4 Andmemudeli seoste loomine Power BI-s

Power Query Editoris loodud tabelite Power BI Desktopi laadimisel ilmnes, et Power BI oli juba automaatselt identifitseerinud dimensioonitabelite võtmed ning nende abil loonud ühendused faktitabeliga. Tehti ülevaatus, et tagada ühenduste korrektne määratlemine.

Tasub märkimist, et andmemudeleid modelleerides ei ole võimalik sätestada Power BI's ühtegi tabeli omadust, mis määraks tabeli dimensiooni või fakti tabeli tüübiks, vaid see määratakse mudeli seoste poolt. Tabeli tüüpi määrab seoste kardinaalsus ja filtreerimise suuna määrab filtri suund. Tavapäraseim suhte kardinaalsus on üks-mitmele või vastupidine mitu-ühele. [13] "Üks" pool on alati dimensiooni tüüpi tabel, kus esineb ainult üks väärtuse eksemplar, samas kui "mitu" pool on alati fakti tüüpi tabel, mille veerus võib olla rohkem kui üks väärtuse eksemplar [13] [14].

Automaatselt loodud ühendus FactTabeli ja DimEriala tabeli vahel on määratud mitu-ühele suhtega ja filtreerimise suunaga fakti tabeli poole (vt Joonis 2.3). See on korrektne, kuna dimensioonitabelist on eemaldatud dubleeritud väärtused ja faktitabel sisaldab kõiki unikaalseid väärtusi.



Joonis 2.3 Seosed andmemudeli tabelite vahel

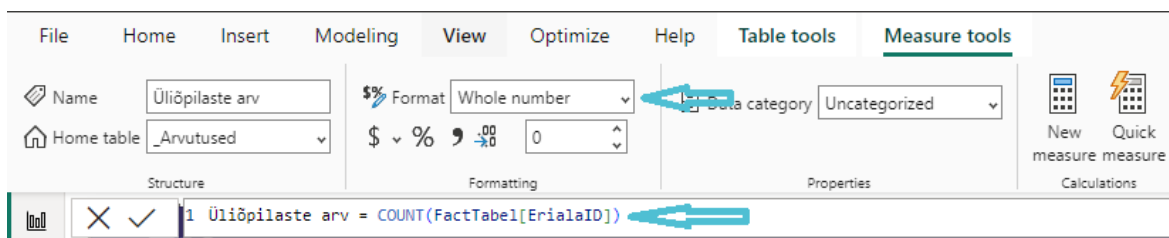
Power BI poolt automaatselt loodud ühendus FactTabeli ja DimÜliõpilane tabeli vahel on üks-ühele seos ja automaatselt on rakendatud kahesuunaline filter (vt Joonis 2.3). Üks-ühele seose korral esineb ühe tabeli veerus ainult üks konkreetne väärtuse eksemplar, ning teises seotud tabelis on samuti ainult üks unikaalne väärtuse eksemplar [13]. Üks-ühele seosed on alati filtreeritud kahesuunaliselt [14]. Kuna dimensioonitabelist DimÜliõpilane ei eemaldatud duplikaate ja see sisaldab kõiki samu Index veeru väärtuseid nagu faktitabelgi, on üks-ühele seos samuti õige.

## 2.5 DAX keeles arvutatud mõõdikute ja veergude lisamine

Power BI Desktop võimaldab spetsiaalselt andmemudelite käitlemiseks loodud valemiväljenduse keeles DAX luua arvutatud veerge, mõõdikuid ja tabelleid andmete dünaamiliselt valimiseks, liitmiseks ja filtreerimiseks [15].

Käesolevas lõputöös TalTech Virumaa kolledži sisseastumisandmete visualiseerimiseks Power BI Desktopis kirjutati DAX-keeles viis mõõdikut. Mõõdikute hoidmiseks loodi uus tabel nimega \_Arvutused.

Järgnevalt on toodud näide mõõdiku "Üliõpilaste arv" loomise protsessist FactTabeli tabeli ErialaID veerus, kus kasutati COUNT() funktsiooni ja valiti täisarvuline vorming (vt Joonis 2.4).



Joonis 2.4 Mõõdik üliõpilaste arvu kokku arvutamiseks

Kokku loodi \_Arvutused tabelisse DAX keelt kasutades 5 mõõdikut, mida näitab alljärgnev Tabel 2.5:

Tabel 2.5 Arvutatud mõõdikud

Mõõdik	Andmetüüp	Kirjeldus	DAX kood
Üliõpilaste arv	Täisarv	Üliõpilaste arv kokku	Üliõpilaste arv = COUNT(FactTabel[ErialaID])
Päevaõppes	Täisarv	Päevaõppes õppivate üliõpilaste arv	Päevaõppes = CALCULATE(COUNT(FactTabel[Index]), DimEriala[Õppevorm] = "päevaõpe")
Sessioonõppes	Täisarv	Sessioonõppes õppivate üliõpilaste arv	Sessioonõppes = CALCULATE(COUNT(FactTabel[Index]), DimEriala[Õppevorm] = "sessioonõpe")
Mees	Täisarv	Meesoost üliõpilaste arv	Mees = CALCULATE(COUNT(FactTabel[Index]), 'DimÜliõpilane'[Sugu]= "mees")
Naine	Täisarv	Naissoost üliõpilaste arv	Naine = CALCULATE(COUNT(FactTabel[Index]), 'DimÜliõpilane'[Sugu]= "naine")

DimÜliõpilane tabelisse lisati Power BI Desktopis DAX keelt kasutades kaks arvutatud

veergu. Loodi veerg Lühike maakond, et kuvada maakonna nimetustest lühikest versiooni ja veerg Vanusevahemik, et oleks võimalik visualiseerida vanusevahemikke viie aasta kaupa. DimÜliõpilane tabelisse loodud arvutatud veergusid näitab Tabel 2.6:

Tabel 2.6 Arvutatud veerud

<b>Veerg</b>	<b>Andmetüüp</b>	<b>Kirjeldus</b>	<b>Dax kood</b>
Lühike maakond	Tekst	Maakonna nimetus ilma sõnata „maakond“	LühikeMaakond = VAR PikkMaakond= 'DimÜliõpilane'[Maakond] VAR TyhikuPositsioon= IFERROR(SEARCH(" ",PikkMaakond),0) RETURN IF(TyhikuPositsioon>0,LEFT(Pikk Maakond,TyhikuPositsioon- 1),PikkMaakond)
Vanusevahemik	Tekst	Üliõpilaste vanusevahemikud grupeeritud viie aasta kaupa	Vanusevahemik = FORMAT(INT('DimÜliõpilane'[Van us] / 5) * 5, "0") & "-" & FORMAT(INT('DimÜliõpilane'[Van us] / 5) * 5 + 5, "0")



### 3. VISUALISEERIMINE

Andmete visualiseerimise peamiseks eesmärgiks on pakkuda kognitiivset tuge teabe mõistmiseks. Graafika võimaldab vaatajal näha suhteid, korrelatsioone, trende ja mustreid, mis vaid numbrite kujul poleks märgatavad. Andmete esitamine visuaalselt lisab meeldejäävust, kiirendab andmete mõistmist, struktureerib infot ja haarab tähelepanu rohkem kui vaid numbrilised loendid [16].

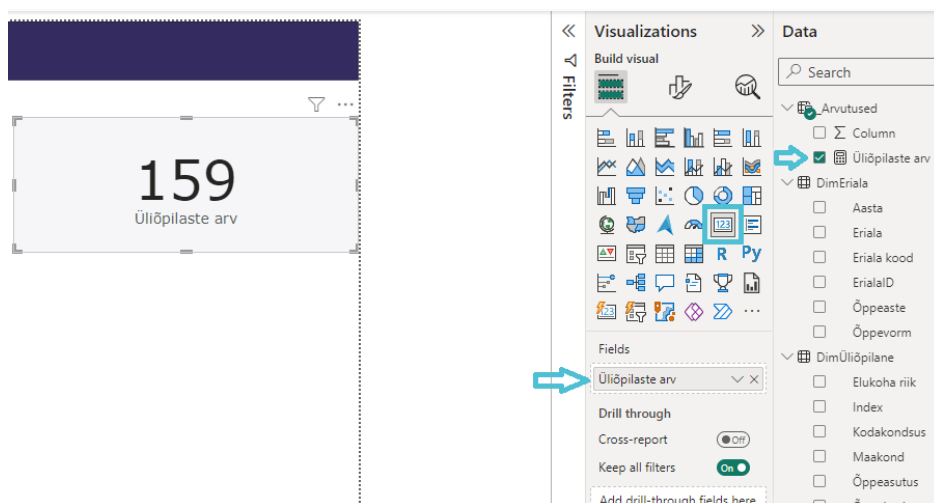
Andmetega töötamisel on oluline teada, millist graafikut valida ja kuidas seda kujundada. Võrdluste tegemiseks sobivad näiteks kõige paremini tulpdiagrammid, kõige efektiivsemad trendide kujutamiseks ajas on joongraafikud ja sektordiagramme kasutatakse terviku osade esitamiseks [16].

Järgnevates alapeatükkides on käsitletud näiteid kõigi visuaali tüüpide kohta, mis aruande koostamiseks loodi. Välja on jäetud filtri visuaali loomine (vt Joonis 4.10), sest seda käsitletakse peatükis 3.2 aruande päise loomise kontekstis (vt Joonis 4.11 ja Joonis 4.12). Eraldi ei ole kirjeldatud ka tulpdiagrammide loomist, sest nende loomiseks kasutatakse rühmitatud tulpdiagrammi ikooni.

Alapeatükid on järjestatud visuaalide aruandes esinemise järjekorras ning kirjeldavad visuaalide loomist ja kujundamist. Kuna visuaalide kujunduse kohandamiseks tehtavaid toiminguid ei korrata, tuleks peatükki suhtuda kui tervikusse.

#### 3.1 Kaardi visuaalide loomine

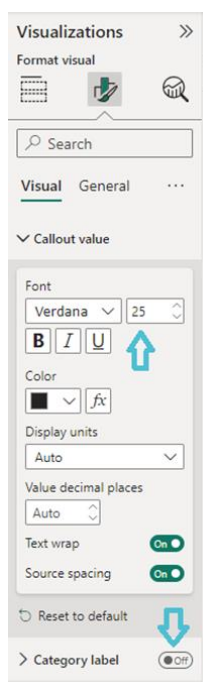
Käesolevas aruandes on kaardi visuaali kasutatud vastuvõetud üliõpilaste arvu visualiseerimiseks. Kaardi visuaalid loodi, valides *Visualizations - Build visual* menüüst *Card* ikoon ja *Data* menüüst *\_Arvutused* tabeli alt *Üliõpilaste arv* (vt Joonis 3.1).



Joonis 3.1 Kaardi visuaali loomine

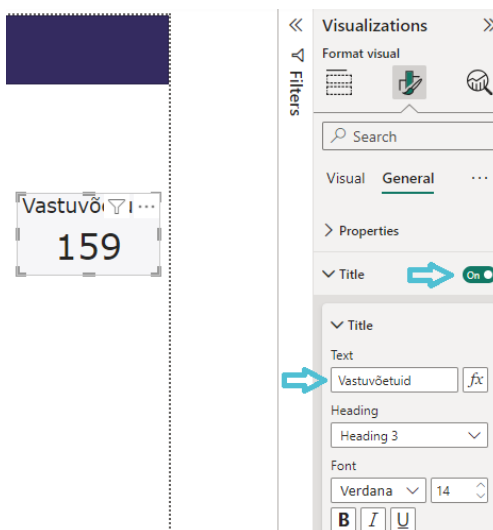
Eemaldati kategooria silt *Visualizations - Format visual - Visual* menüüst ja muudeti

Callout value menüüst väärtuse kirja pildi suurust (vt Joonis 3.2).



Joonis 3.2 Visuaali väärtuse kirja pildi muutmine ja kategooria sildi eemaldamine

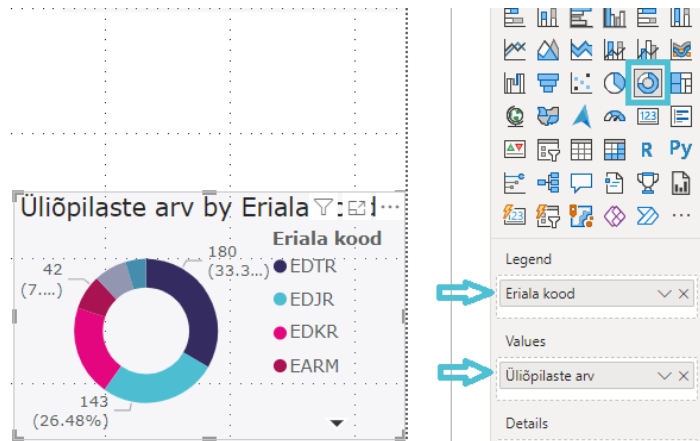
Visuaalide pealkirjad aktiveeriti *Visualizations – Format visual – General – Title* menüüs, mis võimaldas pealkirjade muutmist lahtris *Text* (vt Joonis 3.3).



Joonis 3.3 Visuaali pealkirja muutmine

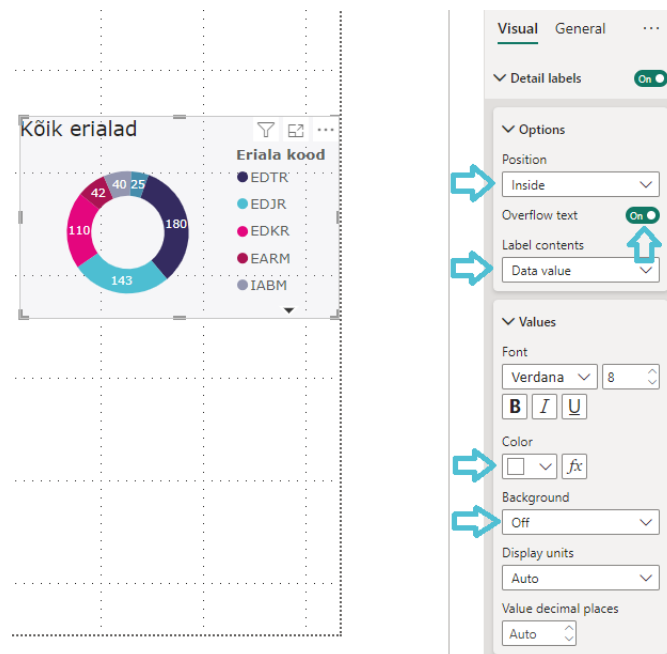
## 3.2 Ringdiagrammide loomine

Sisseastumise statistika visualiseerimisel kasutati ringdiagramme vastuvõetud üliõpilaste arvu jaotuse visualiseerimiseks eriala, soo, vanuse ja kodakondsuse lõikes. Ringdiagrammid loodi, valides *Visualizations – Build visual* menüüst *Donut chart* ikoon ja valides *Legend* ja *Values* lahtritesse *Data* menüüst vastavad väljad (vt Joonis 3.4).



Joonis 3.4 Ringdiagrammi loomine

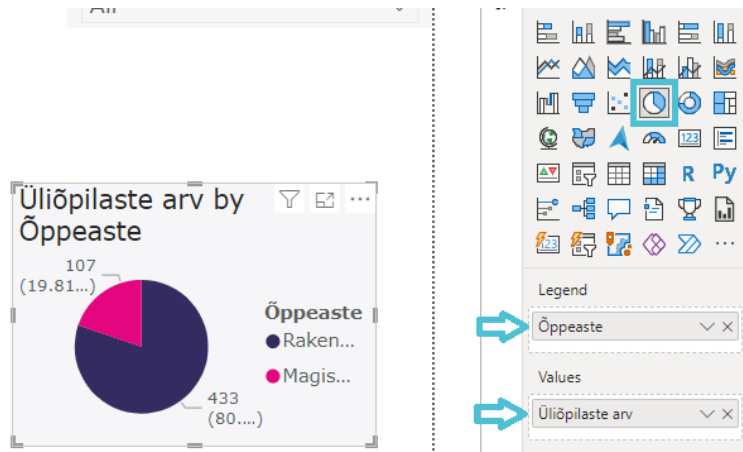
Visuaalide andmesilte võimaldas muuta andmesiltide parameetrite muutmine *Visual - Data labels* menüüs (vt Joonis 3.5).



Joonis 3.5 Andmesiltide muutmine

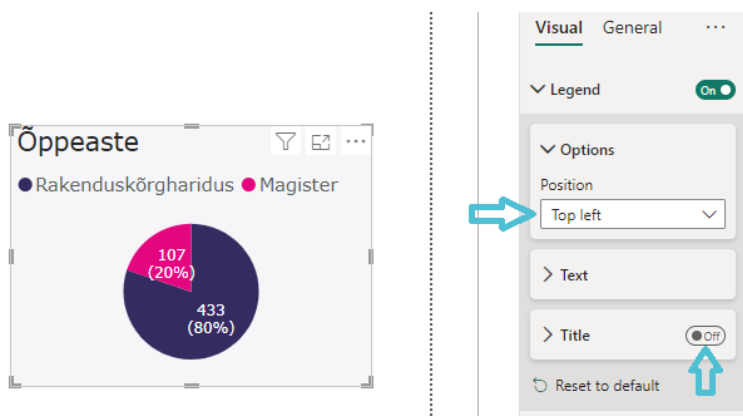
### 3.3 Sektordiagrammide loomine

Sisseastumise statistika visualiseerimisel kasutati sektordiagramme vastuvõetud üliõpilaste arvu jaotuse visualiseerimiseks õppeastme ja õppevormi lõikes. Sektordiagrammid lisati aruandele valides *Vizualizations - Build visual* menüüst *Pie chart* ikoon, ja määrates *Legend* ja *Values* lahtritesse *Data* menüüst vastavad väljad (vt Joonis 3.6).



Joonis 3.6 Pirukadiagrammi loomine

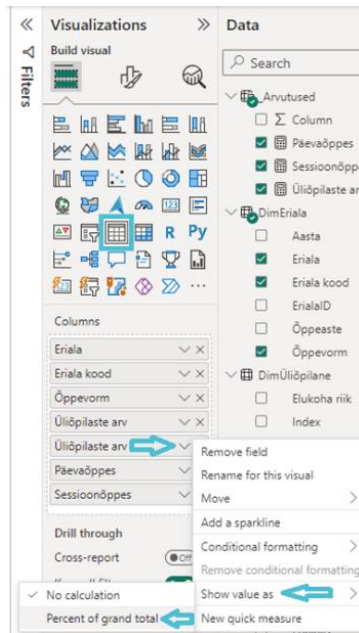
Visuaalide legendi võimaldas muuta *Visual – Legend* menüü (vt Joonis 3.7).



Joonis 3.7 Visuaali legendi muutmise

### 3.4 Tabeli visuaali loomine

Tabeli visuaale kasutati käesolevas lõputöös selleks, et välja tuua informatsiooni üliõpilaste arvu jagunemise kohta erialade, maakondade, kodakondsuse ja varasema õppeasutuse järgi. Tabeli visuaalid loodi valides *Vizualizations - Build visual* menüüst *Table* ikoon ja *Data* menüüst vastavad tabeli veerud. Et kuvada näiteks nii üliõpilaste arvu kui ka üliõpilaste arvu kogusumma protsenti, lohistati hiirega üliõpilaste arvu väärtus lahtrisse *Columns* kahel korral ja valiti rippmenüüst väärtuse näitamine kogusumma protsendina *Show value as – Percent of Grand total* (vt Joonis 3.8).



Joonis 3.8 Tabeli visuaali loomine

Tabelite kujunduseks valiti *Visualizations – Format visual – Visual – Style presets* menüüst *Minimal* ja muudeti tabelite võrestiku värvid vastavalt ülikooli värvigammale (vt Joonis 3.9).

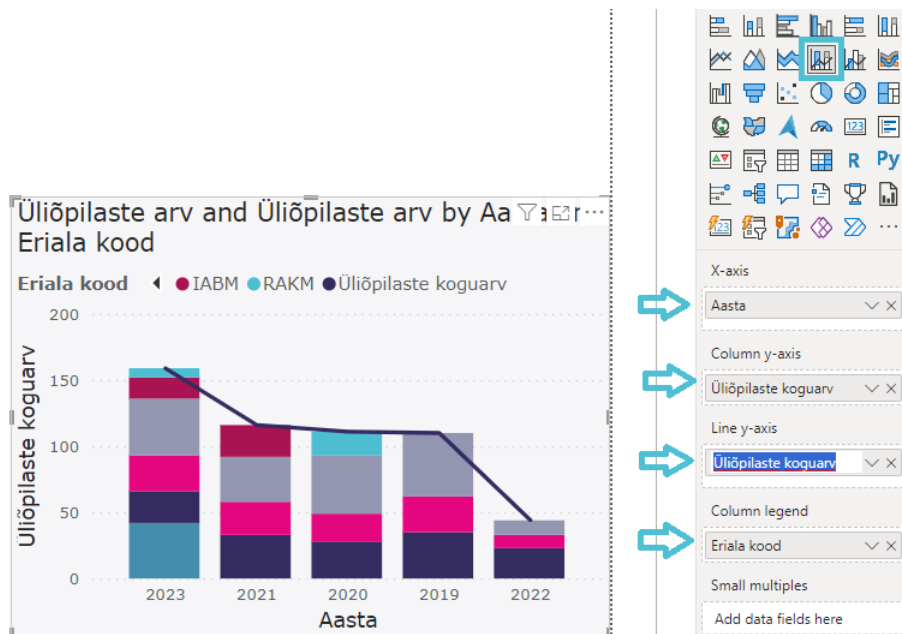
Eriala	Eriala kood	Õppevorm	Üliõpilaste arv	%GT Üliõpilaste arv	Päevaõppes	Sessioonõppes
Äriinfotehnoloogia	IABM	sessioonõpe	40	7.41%		40
Arukad süsteemid ja rakendusinfotehnoloogia	EDTR	päevaõpe	10	1.85%	10	33
Arukad süsteemid ja rakendusinfotehnoloogia	EDTR	sessioonõpe	33	6.11%	10	33
Keemiatehnoloogia	EDKR	päevaõpe	57	10.56%	57	53
Keemiatehnoloogia	EDKR	sessioonõpe	53	9.81%	57	53
Kütuste keemia ja tehnoloogia	RAKM	sessioonõpe	25	4.63%		25
<b>Total</b>			<b>540</b>	<b>100.00%</b>	<b>163</b>	<b>377</b>

Joonis 3.9 Tabeli kujunduse muutmine

### 3.5 Joon- ja virnastatud tulpdiagrammide loomine

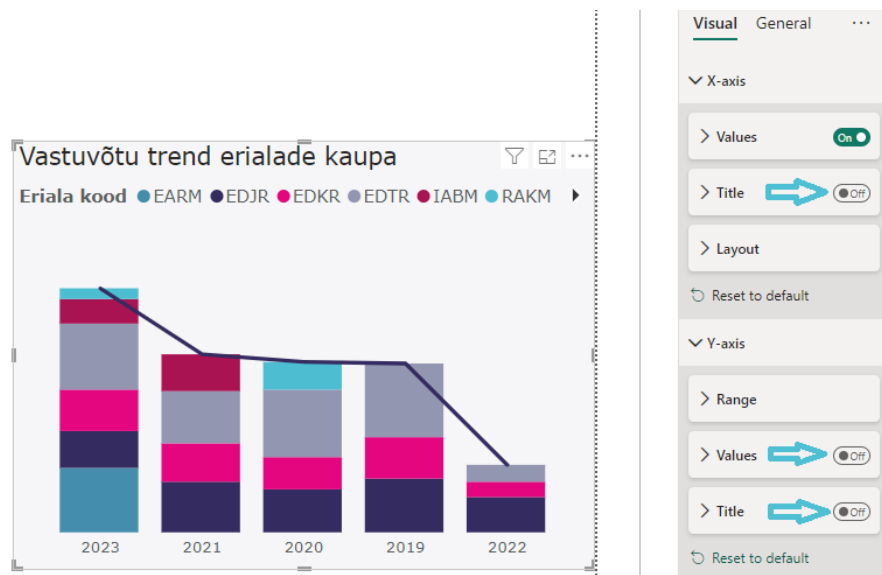
Joon- ja virnastatud tulpdiagramme kasutati aruandes aastate 2019–2023 vastuvõtu trendide visualiseerimiseks eriala, õppeastme ja õppevormi lõikes. Joon- ja virnastatud tulpdiagrammid loodi valides *Visualizations – Build visual* menüüst *Line and stacked*

column chart ikoon ja paigutades *Data* menüüst *X-axis*, *Column y-axis*, *Line y-axis* ja *Column legend* lahtritesse vastavad väärtused. Väärtuste nimetusi visuaalil võimaldas muuta kaks hiireklikki nimetusel (vt Joonis 3.10).



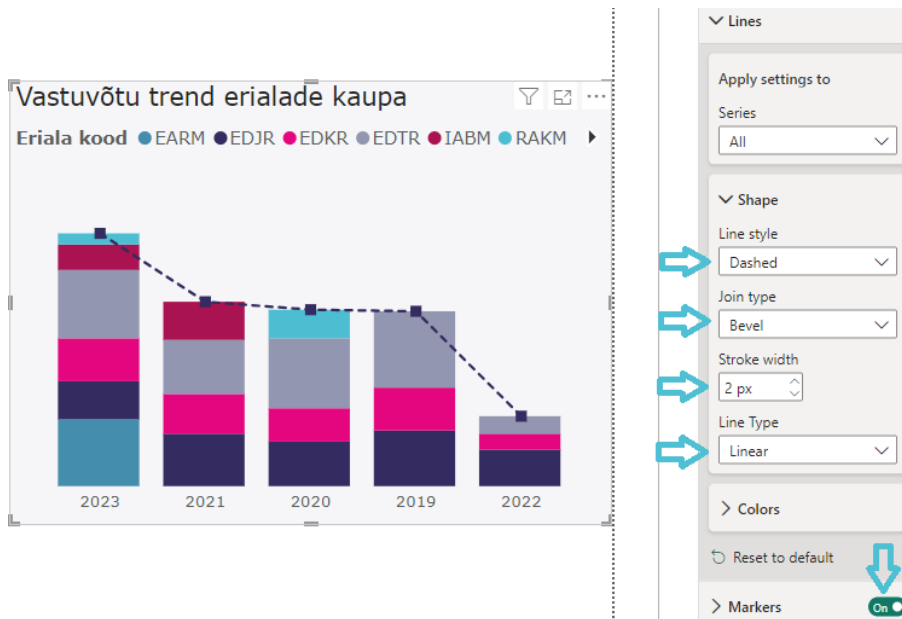
Joonis 3.10 Joon- ja virnastatud tulpdiagrammi loomine

Selleks et visuaali disaini puhtamaks muuta, eemaldati *Visual* menüüs x-teljelt ja y-teljelt pealkirju ja väärtuseid (vt Joonis 3.11).



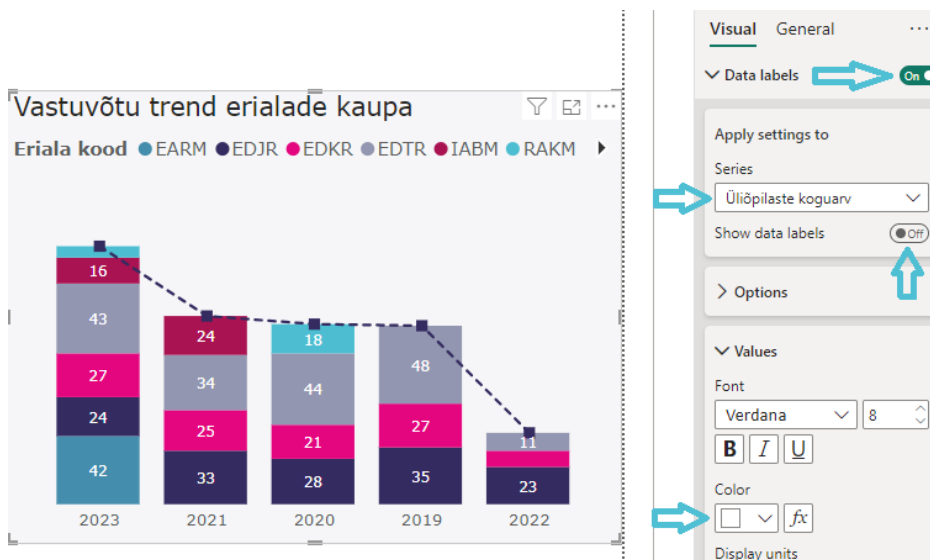
Joonis 3.11 x- ja y-telje muutmine

Joone disaini võimaldas muuta *Visuals* – *Lines* menüü, kus oli võimalik muuta nii joone kuju kui ka lisada sellele markereid (vt Joonis 3.12).



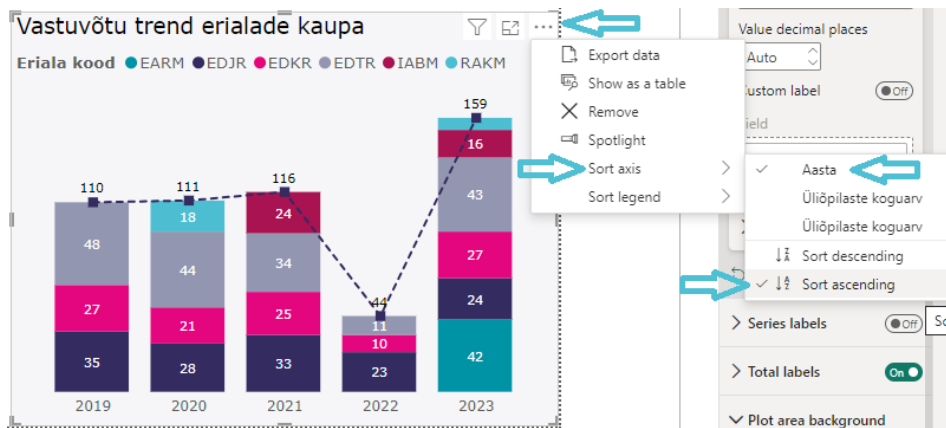
Joonis 3.12 Joone disaini muutmine

Joon- ja virnastatud tulpdiaagrammidele andmesiltide lisamine käis sarnaselt, kui ringdiagrammidele (vt Joonis 3.5), kuid erinevuseks oli see, et lisaks oli võimalik disainida andmesilte vastavalt visuaals kasutatud väärtustele. Antud näites lisati andmesildid kõigile väärtustele ja seejärel lülitati y-telje joondiagrammil andmesildid välja (vt Joonis 3.13).



Joonis 3.13 Joondiagrammi andmesiltide muutmine

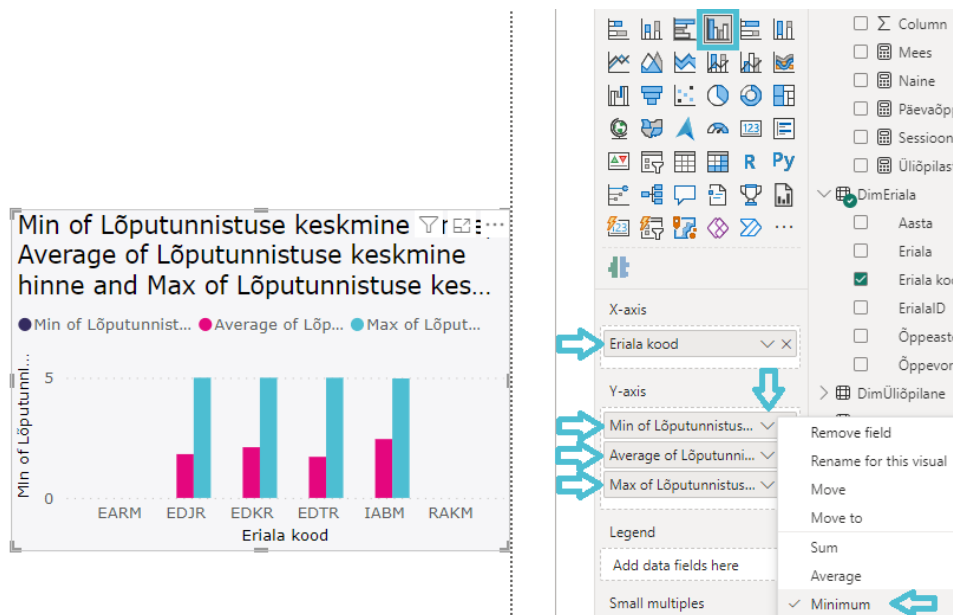
Visuaalide sorteerimiseks valiti visuaali paremas nurgas kolmele täpile vajutades ilmuvast rippmenüüst *Sort axis*, mis võimaldas valida sorteerimise aluseks väärtuse ja sorteerida kasvavas järjekorras, valides *Sort ascending*, või kahanevas järjekorras, valides *Sort descending* (vt Joonis 3.14).



Joonis 3.14 Visuaali sorteerimine

### 3.6 Rühmitatud tulpdiagrammide loomine

Rühmitatud tulpdiagramme kasutati aruandes selleks, et visualiseerida eriala lõikes üliõpilaste arvu jagunemist õppevormi, vanuse, tunnustuse ja keskmise hinne alusel. Rühmitatud tulpdiagrammid loodi, valides *Vizualizations - Build visual* menüüst *Clustered column chart* ikoon ja paigutades *Data* menüüst vastavad väärtused x-telje ja y-telje või legendi lahtritesse. Selleks et näidata ühe väärtuse miinimum-, keskmist ja maksimaalset väärtust, lohistati sama väärtus *Data* menüüst y-telje lahtrisse kolmel korral ja valiti väärtuse rippmenüüst vastavalt väärtuse kuvamine *Minimum*, *Maximum* ja *Average* järgi (vt Joonis 3.15).

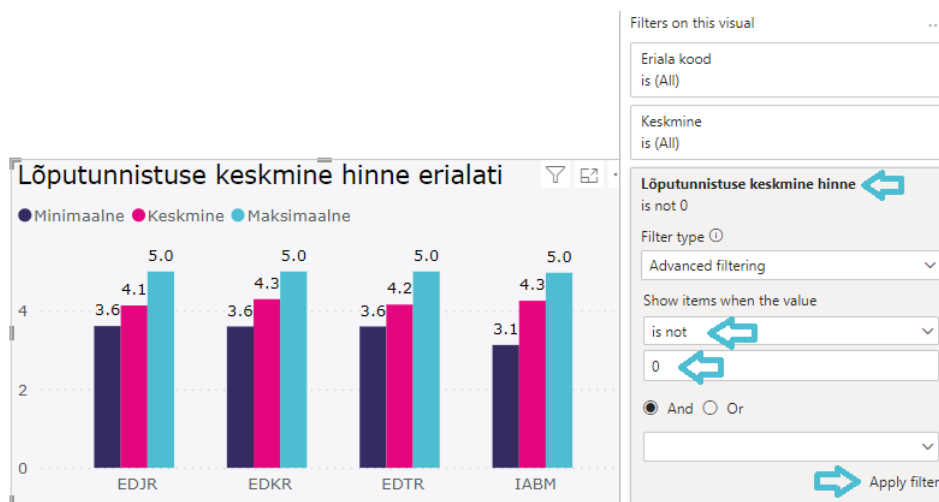


Joonis 3.15 Rühmitatud tulpdiagrammi loomine

Visuaalil või leheküljel kuvatavaid väärtusi võimaldas filtreerida *Filters* paneel. Selleks et rühmitatud tulpdiagramm „Lõputunnistuse keskmine hinne erialati“ ei kuvaks



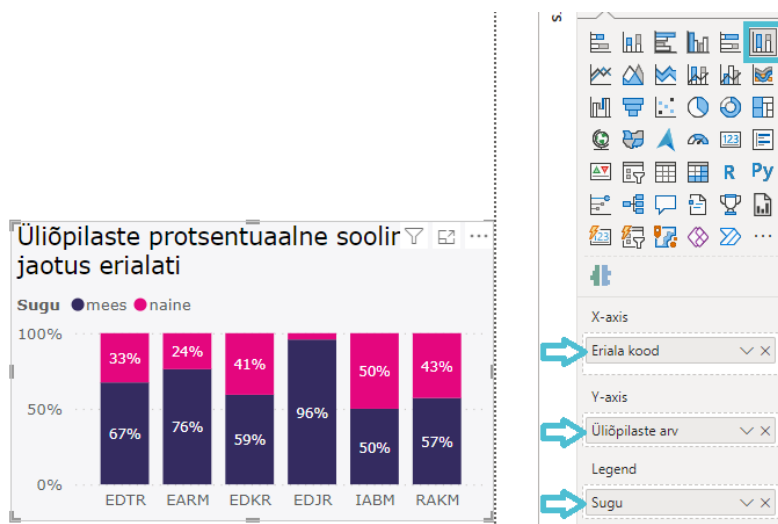
nullväärtuseid, rakendati visuaalile filter, mis välistas keskmise hinne väärtuse 0 (vt Joonis 3.16).



Joonis 3.16 Väärtust välistava filtri lisamine visuaalile

### 3.7 100% virnastatud tulpdiagrammi loomine

100% virnastatud tulpdiagrammi abil visualiseeriti aruandes üliõpilaste arvu protsentuaalset jagunemist soolise jaotuse järgi erialati. 100% virnastatud tulpdiagramm loodi, valides *Vizualizations - Build visual* menüüst *100% Stacked column chart* ikoon ja paigutades *Data* menüüst vastavad väärtused x-telje, y-telje ja legendi lahtritesse. (vt Joonis 3.17).

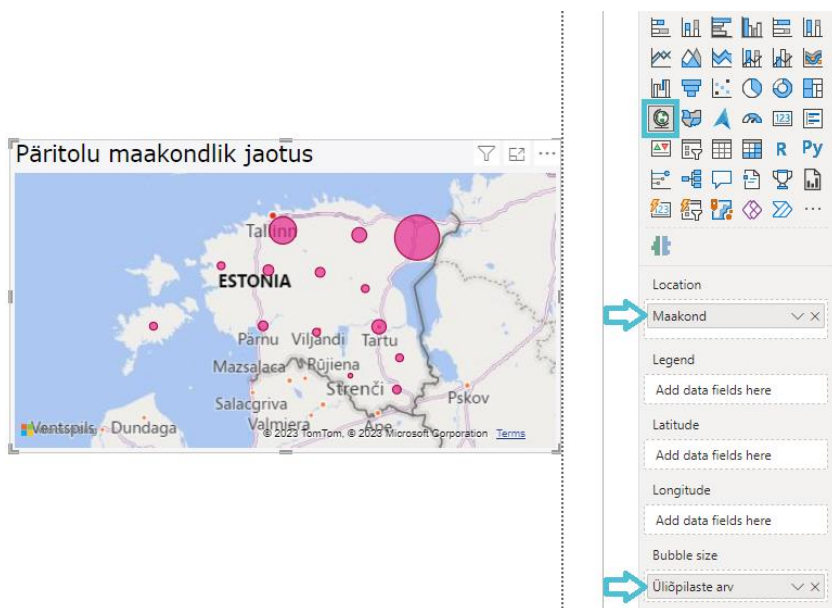


Joonis 3.17 100% virnastatud tulpdiagrammi loomine

### 3.8 Geograafilise kaardi visuaali loomine

Geograafilise kaardi visuaali abil kujutati käesolevas lõputöös üliõpilaste arvu maakondlikku jaotust kaardil. Selleks valiti *Vizualizations - Build visual* menüüst *Map*

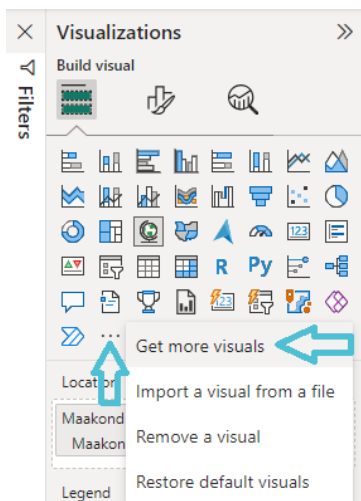
ikoon, määrati *Location* lahtris asukohaks maakonna veerg, ning mulli suuruseks *Bubble size* lahtris üliõpilaste arv (vt Joonis 3.18).



Joonis 3.18 Geograafilise kaardi visuaali loomine

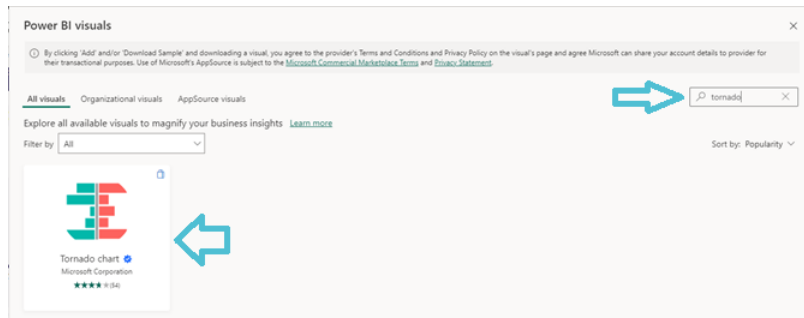
### 3.9 Tornaadodiagrammi loomine

Tornaadodiagrammi abil visualiseeriti aruandes üliõpilaste arvu jagunemist maakondade vahel õppevormi alusel. Kuna Power BI Desktop tarkvara ei sisalda algselt tornaadodiagrammi, kuid võimaldab erinevaid diagramme programmi juurde lisada, siis tehti seda, valides *Build visual* ikoonide menüüst kolme punkti vajutamisel avanevast rippmenüüst *Get more visuals* (vt Joonis 3.19).



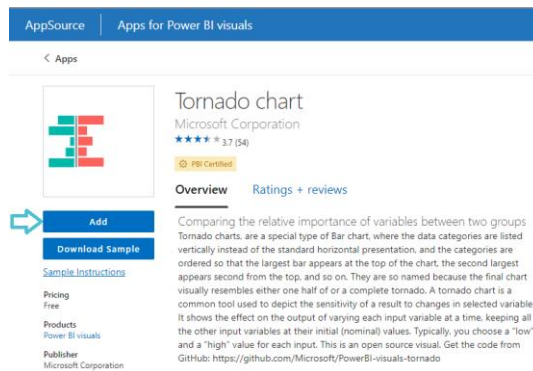
Joonis 3.19 Visuaalide juurde lisamine

Avanevast Power Bi visuaalide loendist otsiti soovitud tornaadodiagramm (vt Joonis 3.20).



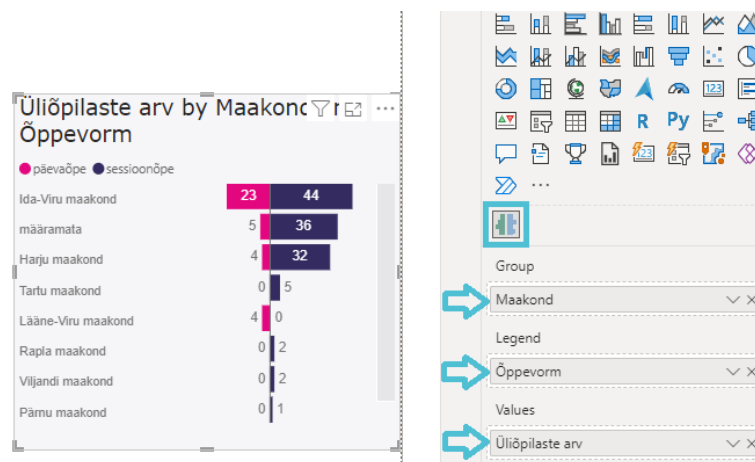
Joonis 3.20 Power BI visuaalide otsimine

Sellele järgnevalt avanes AppSource aken, millest leidis visuaali kohta rohkem informatsiooni. Soovitud visuaali Power BI Desktopi lisamiseks tuli vajutada nupule Add (vt Joonis 3.21).



Joonis 3.21 Rakendused Power BI visuaalidele

Tornaadodiagrammi AppSource kaudu lisamise järgselt ilmus *Build visual* menüüsse tornaadodiagrammi ikoon, mis võimaldas visualiseerida päeva- ja sessioonõppe üliõpilaste arvu jagunemist maakonniti, valides grupeerimiseks maakonna, legendiks õppevormi ja väärtusteks üliõpilaste arvu (vt Joonis 3.22).



Joonis 3.22 Tornaadodiagrammi loomine

## 4. ARUANDE LOOMINE

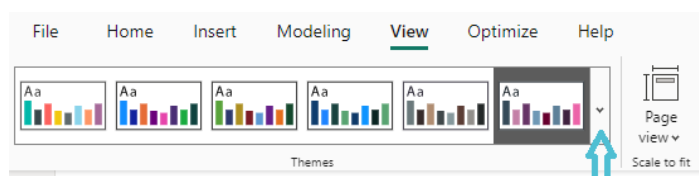
Käesoleva lõputöö põhieesmärgina loodi Power BI desktopis kaheksaleheküljeline aruanne "TalTech Virumaa kolledži sisseastumisandmed" Virumaa kolledži sisseastumise statistika visualiseerimiseks. Igale aruande leheküljele lisati päis, mille kujundamiseks kasutati ristkülikukujulist kujundit, millele lisati logo pilt ja navigeerimisriba. Lisaks loodi päise juurde aasta filtri visuaal. Aruande kujundamisel lähtuti Tallinna Tehnikaülikooli Stiiliraamatus välja toodud brändi eeskirjadest [17].

Tabel 4.1 TalTech Virumaa kolledži vastuvõtustatistika aruande lehekülgede jaotus

Lehekülje nimetus	Lehekülje sisu
Vastuvõtt	Üldine vastuvõtustatistika: vastuvõetud üliõpilaste arv ja selle jaotus eriala, õppeastme ja õppevormi lõikes.
Trendid	Vasuvõtu trendid läbi aastate eriala, õppeastme ja õppevormi lõikes.
Rakendusõpe	Rakenduskõrghariduse erialadele vastuvõtu statistika: vastuvõetud üliõpilaste arv ja selle jaotus eriala, õppeastme ning õppevormi lõikes ning vasuvõtu trendid eriala ja õppevormi lõikes.
Magister	Magistrierialadele vastuvõtu statistika: vastuvõetud üliõpilaste arv ja selle jaotus erialade lõikes ning vasuvõtu trend erialade lõikes.
Sugu	Vastuvõtustatistika üliõpilaste soolise jaotuse järgi, mis on filtreeritav õppeastme ja õppevormi alusel.
Vanus	Vastuvõetud üliõpilaste arvu vanuseline jaotus minimaalse, keskmise ja maksimaalse vanuse lõikes erialati.
Päritolu	Vastuvõetud üliõpilaste kodakondsuse ja maakondliku päritolu statistika, mis on filtreeritav eriala järgi.
Õppeasutus	Vastuvõetud üliõpilaste eelneva õppeasutuse, tunnustuse ja keskmise hinde statistika, mis on filtreeritav eriala alusel.

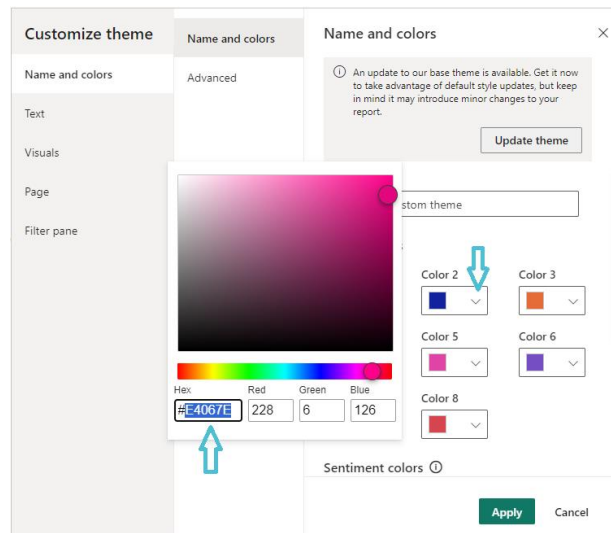
### 4.1 Aruande kujundus

Tallinna Tehnikaülikoolil on välja töötatud oma bränd ning nõuded, mistõttu oli oluline tagada aruande visuaalse stiili vastavus Tallinna Tehnikaülikooli Stiiliraamatule [17]. Enne Power BI aruande koostamist seadistati kujunduse mall vastavalt ülikooli brändinõuetele. Power BI Desktopis avati *View* menüüst *Themes* rippmenüü (vt Joonis 4.1), millest valiti kujunduse kohandamiseks *Customize current theme*.



Joonis 4.1 Kujunduse rippmenüü

*Name and colors* sektsioon võimaldas aktsendivärvid ükshaaval muuta, kopeerides stiiliraamatust õiged HEX-värvikoodid (vt Joonis 4.2).














Joonis 4.2 Värvide muutmine

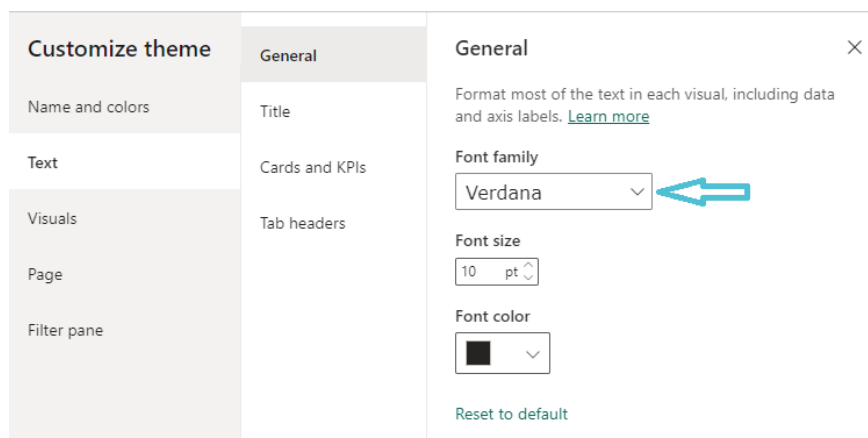
Tuleb märkida, et värvide esinemissagedus aruande kujunduses vähenes vastavalt värvinumbrite järjekorrale. Kuna helehall, must ja valge ei sobinud aktsentvärvideks, siis otsustati need asendada ülikooli värvide gradienttoonidega, kasutades selleks spetsiaalset gradiendi generaatorit [18].

Allpool esitatakse Tabel 4.2, milles on toodud kujunduses kasutatud värvitoonid, nende nimetused, HEX-koodid ja nende kasutusala.

Tabel 4.2 Kujunduses kasutatud värvid

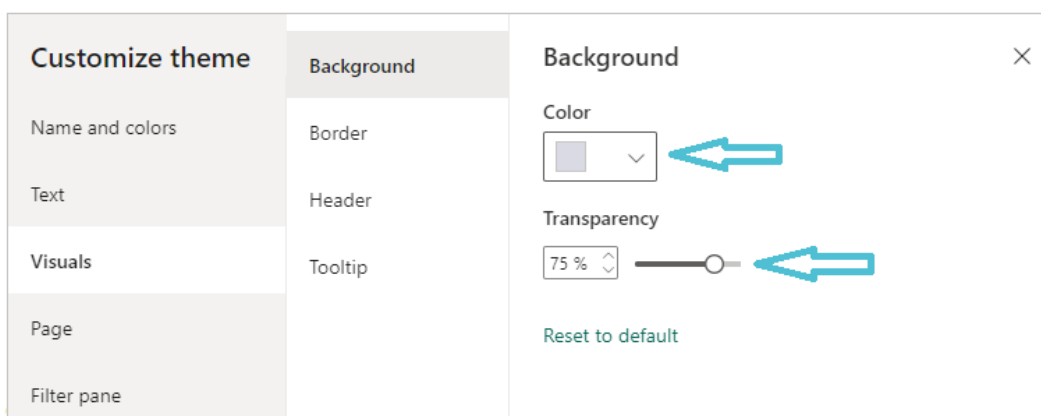
Värv	Nimetus	HEX-kood	Kasutus
	Tumesinine	#342b60	Aktsendivärv 1
	Fuksiapunane	#e4067e	Aktsendivärv 2
	Helesinine	#4dbed2	Aktsendivärv 3
	Kirsipunane	#aa1352	Aktsendivärv 4
	Terashall	#9396b0	Aktsendivärv 5
	Sinine (helesinise ja musta gradientvärv)	#0093a6	Aktsendivärv 6
	Pruun (kirsipunase ja musta gradientvärv)	#6d0025	Aktsendivärv 7
	Lilla (tumesinise ja fuksiapunase gradientvärv)	#6a2f7a	Aktsendivärv 8
	Helehall	#000000	Visuaali tausta värv
	Valge	#FFFFFF	Lehekülje tausta värv
	Must	#dadae4	Teksti värv

Kuna Power BI-s ei olnud saadaval TalTech-i brändi fonti Proxima Nova, siis kasutati alternatiivina soovitatud Verdana fonti. Selleks valiti *Text - General, Title, Cards and KPIs* ja *Tab headers* sektsioonist *Font Family* Verdana (vt Joonis 4.3).



Joonis 4.3 Fondi muutmise

Visuaalidele taustaks määrati *Visuals - Background* sektsioonist TalTechi brändi helehall värvitoon ja sätestati läbipaistvus 75% peale (vt Joonis 4.4).

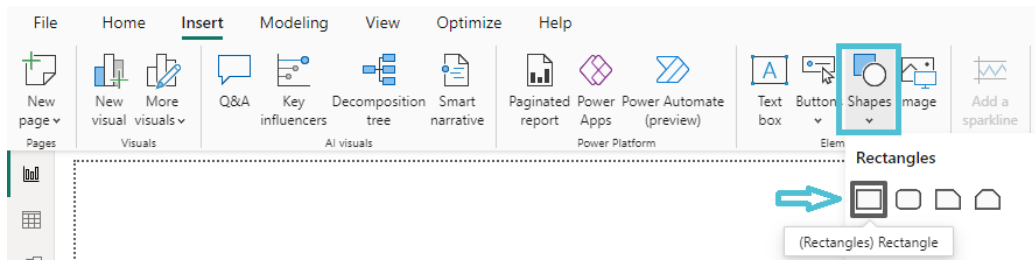


Joonis 4.4 Visuaalide taust

## 4.2 Aruande päise loomine

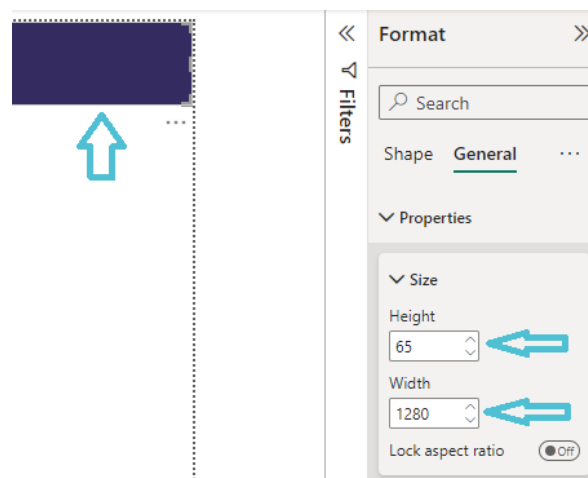
Aruandele professionaalsema ja korrastatuma visuaalse ilme tagamiseks loodi päis, mis võimaldab kasutajal kiiret ja selget navigeerimist selle sisus. Päis mängib olulist rolli aruande konteksti mõistmisel, kuna see sisaldab Tallinna Tehnikaülikooli logo ja aruande lehekülgede pealkirju. Aasta filtri integreerimine päisega aitab selgitada kasutajatele, et see mõjutab terve lehekülje sisu.

Aruande päise kujundamiseks valiti *Insert* menüüst *Shapes - Rectangle* (vt Joonis 4.5).



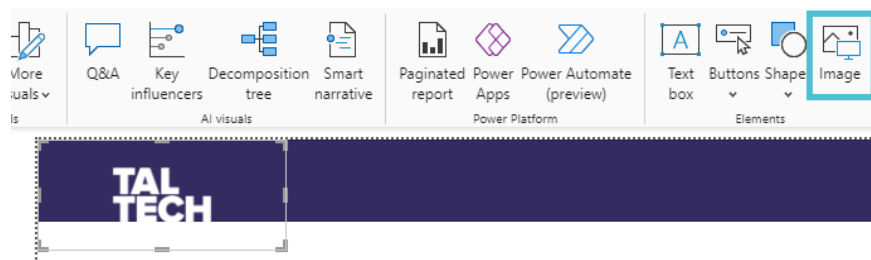
Joonis 4.5 Kujundi lisamine

Ilmunud ristkülikukujulise elemendi suurust reguleeriti aruande lehekülje ühest horisontaalsest äärest teise, määrates kõrguseks *Format – General – Properties* menüüst väärtuse 65 (vt Joonis 4.6).



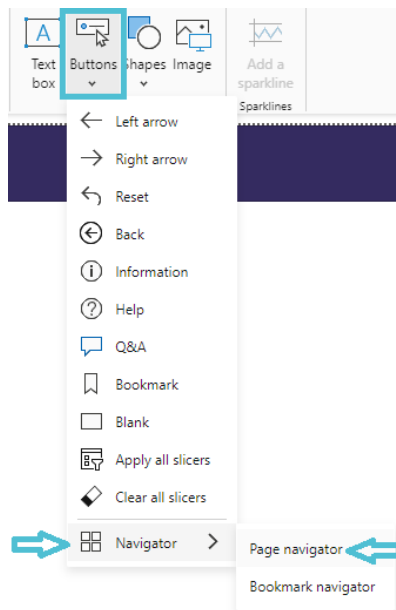
Joonis 4.6 Suuruse määramine

Järgmiseks lisati pilt TalTechi logoga, mis oli eelnevalt arvutisse salvestatud. Selleks valiti *Insert* menüüst *Image* ja paigutati pilt päise vasakusse äärde (vt Joonis 4.7).



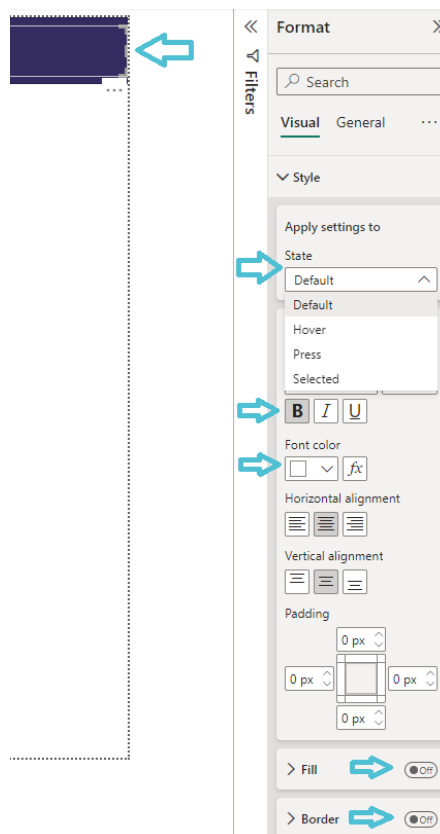
Joonis 4.7 Pildi lisamine

Lehekülje navigaatori lisamiseks päisele valiti *Insert* menüüst *Buttons – Navigator – Page navigator* (vt Joonis 4.8).



Joonis 4.8 Lehekülje navigaatori lisamine

Navigaatori paigutati päise keskele ning kujundust kohandati *Format – Visual – Style* menüüst (vt Joonis 4.9).



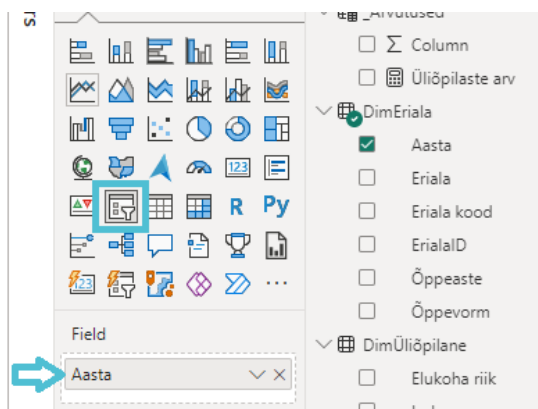
Joonis 4.9 Navigaatori kujundus

Oluline on märkida, et navigaatoril on neli olekut, ning kujunduse muudatused rakendati igas olekus. Vaikimisi olekus *Default* muudeti teksti värv valgeks ja tekstipaksuseks



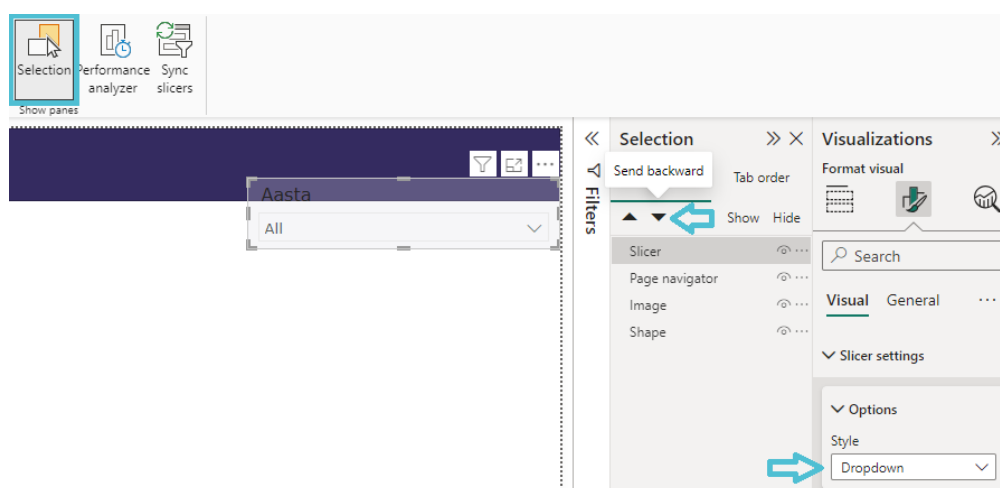
valiti *Bold*. Deaktiveeriti navigaatori sisu ja ääred, lülitades parameetrid *Fill* ja *Border* asendisse *Off*. Hiirega teksti kohal oleku olekus *Hover* muudeti teksti värv helehalliks, tekstile vajutamise olekus *Press* terashalliks ning aktiivse lehekülje olekus *Selected* helesiniseks.

Järgmise sammuna loodi päise juurde filtri visuaal võimaldamaks informatsiooni aruande lehekülgedel filtreerida vastavalt aastale. Selleks valiti *Visualizations – Build visual* menüüst filtri ikoon *Slicer*. DimEriala tabelist kasutati aasta välja (vt Joonis 4.10).



Joonis 4.10 Filtri visuaali loomine

Seejärel kohandati filtri visuaali välimust, muutes selle rippmenüüks. Selleks valiti menüüst *Visualizations - Format visual - Visual – Slicer settings – Options* filtri tüübiks *Dropdown*. Filtri visuaal paigutati osaliselt päise peale ja seejärel liigutati päise taha peitu, aktiveerides *Format* menüüst *Selection* külgmenüü, kus oli võimalik kihtide järjekorda muuta ja filtri kiht paigutada teiste päiselementide selektsioonikihtide taha (vt Joonis 4.11)



Joonis 4.11 Rippmenüüks kohandamine ja selektsioonikihtide järjekorra muutmine

Järgnevalt on esitatud pilt aruande päisest valmis kujul (vt Joonis 4.12). Päise

navigatsioonimenüü hõlmab logo ja kaheksat aruande lehekülje pealkirja, võimaldades kasutajal aruande lehekülgede vahel mugavalt liikuda. Päisele on lisatud aasta filtri rippmenüü, mis võimaldab visuaale vastavalt valitud aastale aruande lehekülgedel muuta.



Joonis 4.12 Aruande päis

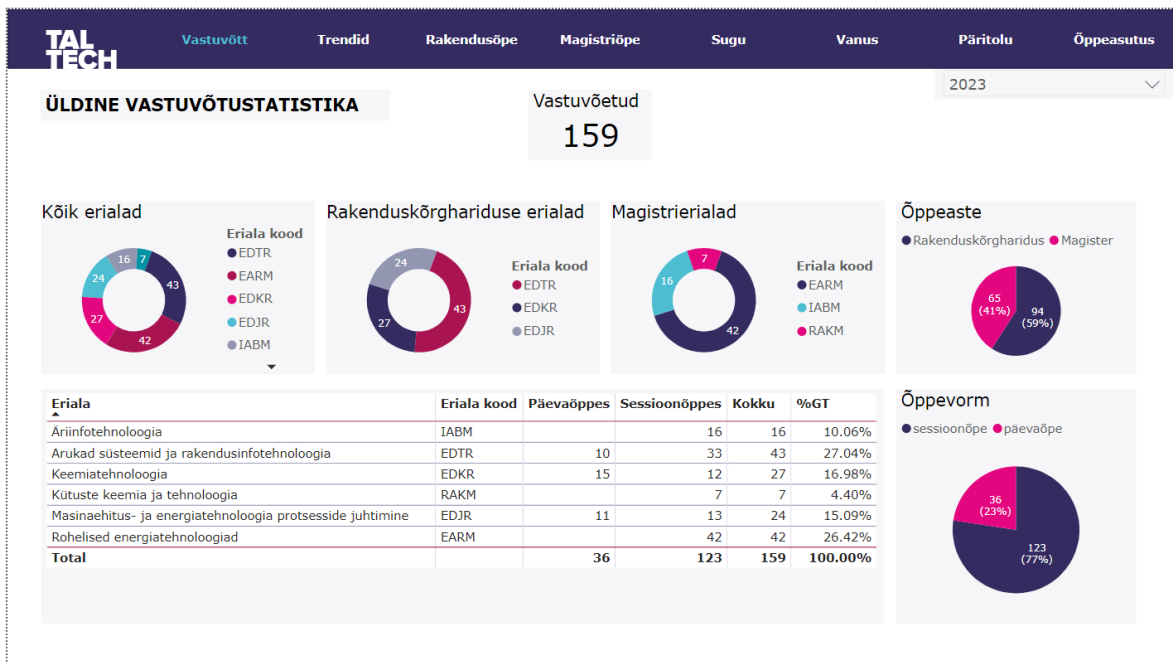
## 4.3 Aruande leheküljed

Aruandele "TalTech Virumaa kolledži sisseastumisandmed" loodi kokku kaheksa aruande lehekülge. Järgnevates alapeatükkides antakse ülevaade kõigist aruande lehekülgedest ja neile loodud visuaalidest.

### 4.3.1 Üldine vastuvõtustatistika

TalTech Virumaa kolledži üldise vastuvõtustatistika visualiseerimiseks loodi aruande lehekülg "Vastuvõtt" (vt Joonis 4.13) pealkirjaga „Üldine vastuvõtustatistika“, kuhu loodi järgmised visuaalid:

- **Kaardi visuaal** „Vastuvõetud“, et kuvada vastuvõetud üliõpilaste koguarvu.
- **Ringdiagramm** "Kõik erialad", et visualiseerida vastuvõetud üliõpilaste arvu erialade lõikes kõigil erialadel.
- **Ringdiagramm** „Rakenduskõrghariduse erialad“, mis visualiseerib vastuvõetud üliõpilaste arvu rakenduskõrghariduse erialade lõikes.
- **Ringdiagramm** "Magistrierialad", mis näitab vastuvõetud üliõpilaste arvu magistrierialadel.
- **Sektordiagramm** "Õppeaste", et visualiseerida rakenduskõrghariduse ja magistriõppesse astunud üliõpilaste arvu numbriliselt ja protsentuaalselt.
- **Sektordiagramm** "Õppevorm", et näidata sessioon- ja päevaõppesse astunud üliõpilaste arvu numbriliselt ja protsentuaalselt.
- **Tabel**, kus on esitatud eriala nimetus, eriala kood, päeva- ja sessioonõppesse sisse astunud üliõpilaste arv, üliõpilaste koguarv ning üliõpilaste arvu kogusumma protsentuaalselt.

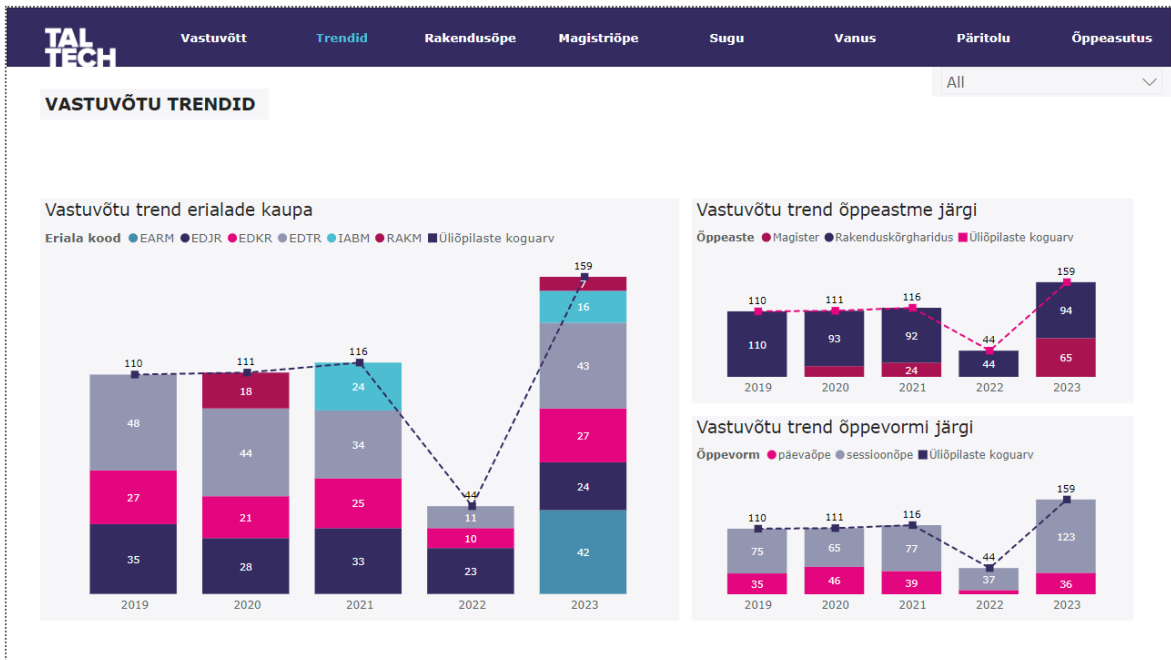


Joonis 4.13 Aruande lehekülj "Vastuvõtt"

### 4.3.2 Vastuvõtu trendid

Aruande leheküljel "Trendid" (vt Joonis 4.14), mis kannab pealkirja „Vastuvõtu trendid“ on kujutatud TalTech Virumaa kolledži vastuvõtu trende aastate lõikes järgmiste visuaalide abil:

- **Joon- ja virnastatud tulpdiagramm** "Vastuvõtu trend erialade kaupa", mis visualiseerib vastuvõetud üliõpilaste arvu trendi läbi aastate kõigi erialade lõikes võrreldes üliõpilaste koguarvuga.
- **Joon- ja virnastatud tulpdiagramm** "Vastuvõtu trend õppeastme järgi", mis näitab magistri- ja rakenduskõrgharidusõppesse vastu võetud üliõpilaste trendi läbi aastate võrreldes üliõpilaste koguarvuga.
- **Joon- ja virnastatud tulpdiagramm** "Vastuvõtu trend õppevormi järgi", näidates päeva- ja sessioonõppesse vastuvõetud üliõpilaste trendi läbi aastate võrreldes üliõpilaste koguarvuga.



Joonis 4.14 Aruande lehekülg "Trendid"

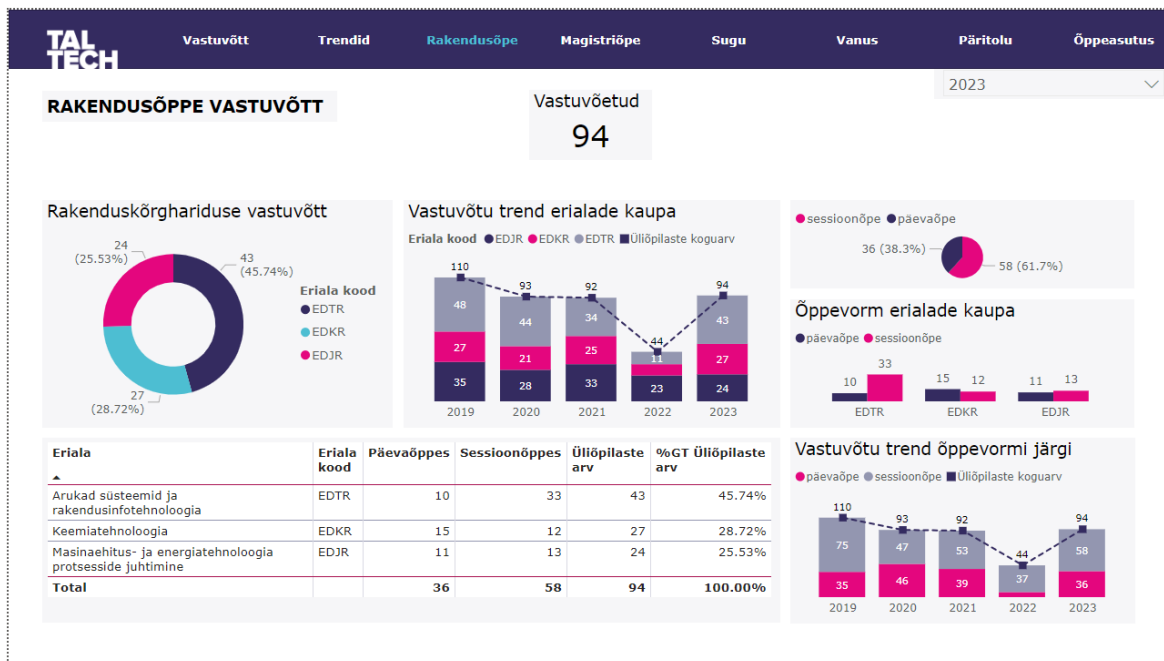
### 4.3.3 Rakendusõppe vastuvõtt

TalTech Virumaa kolledži rakenduskõrghariduse vastuvõtustatistika visualiseerimiseks loodi aruande lehekülg „Rakendusõpe“ (vt Joonis 4.15) pealkirjaga „Rakendusõppe vastuvõtt“. Lehekülg filtreeriti rakenduskõrghariduse õppevormi järgi ja sellele loodi järgmised visuaalid:

- **Kaardi visuaal** „Vastuvõetud“, mis kuvab rakenduskõrghariduse erialadele vastuvõetud üliõpilaste arvu.
- **Ringdiagramm** „Rakenduskõrghariduse vastuvõtt“, mis visualiseerib rakenduskõrghariduse erialadele vastu võetud üliõpilaste arvu nii numbriliselt kui ka protsentuaalselt.
- **Sektordiagramm**, mis näitab rakenduskõrghariduse erialadele vastu võetud üliõpilaste arvu päeva- ja sessioonõppe lõikes nii numbriliselt kui ka protsentuaalselt.
- **Joon- ja virnastatud tulpdiagramm** „Vastuvõtu trend erialade kaupa“, et illustreerida vastuvõetud üliõpilaste arvu trendi läbi aastate kõikidel rakenduskõrghariduse erialadel võrreldes üliõpilaste koguarvuga.
- **Joon- ja virnastatud tulpdiagramm** „Vastuvõtu trend õppevormi järgi“, et kajastada rakenduskõrghariduse erialadele vastuvõetud üliõpilaste arvu trendi päeva- ja sessioonõppes läbi aastate võrreldes üliõpilaste koguarvuga.
- **Rühmitatud tulpdiagramm** „Õppevorm erialade kaupa“, mis näitab päeva ja

sessioonõppesse astunud üliõpilaste arvu rakenduskõrghariduse erialade kaupa.

- **Tabel**, kus on esitatud eriala nimetus, eriala kood, päeva- ja sessioonõppesse astunud üliõpilaste arv, üliõpilaste koguarv ning üliõpilaste arvu kogusumma protsentuaalselt.

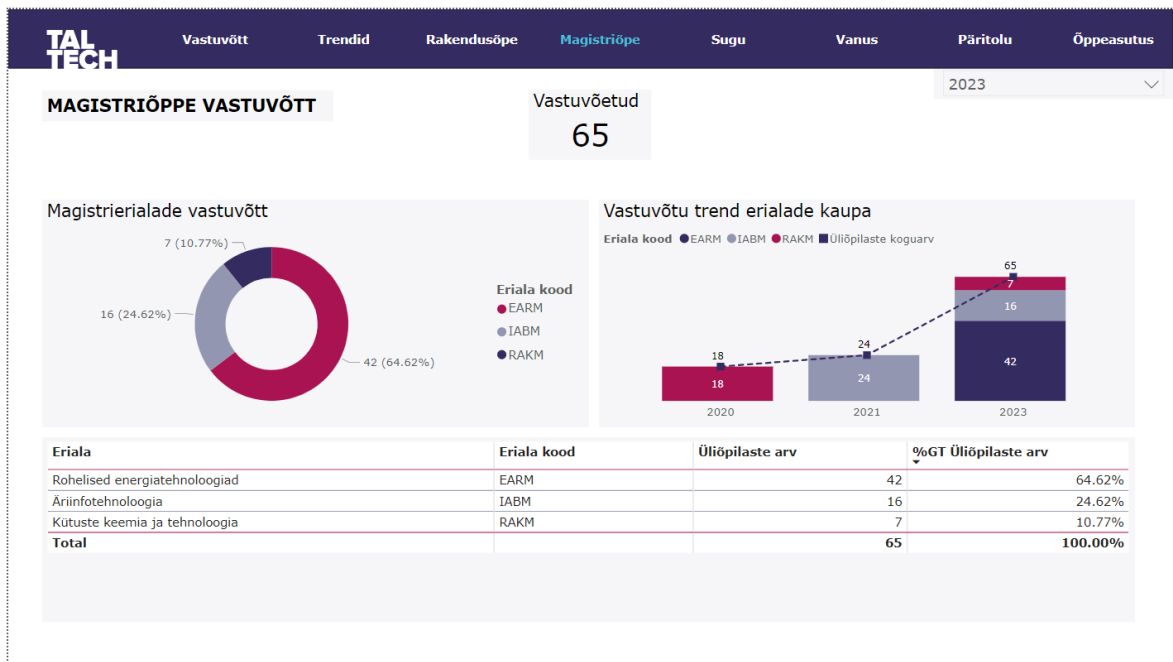


Joonis 4.15 Aruande lehekülg „Rakendusõpe“

### 4.3.4 Magistriõppe vastuvõtt

Selleks et visualiseerida TalTech Virumaa kolledži magistrerialadele vastuvõtu statistikat, loodi aruande lehekülg „Magistriõpe“ (vt Joonis 4.16) pealkirjaga „Magistriõppe vastuvõtt“, millele lisati filtreerimine magistriõppe järgi ja kuhu loodi järgmised visuaalid:

- **Kaardi visuaal** „Vastuvõetud“, mis kuvab magistrerialadele vastu võetud üliõpilaste arvu.
- **Ringdiagramm** „Magistrerialade vastuvõtt“, visualiseerimaks magistrerialadele vastu võetud üliõpilaste arvu nii numbriliselt kui ka protsentuaalselt.
- **Tabel**, kus on esitatud eriala nimetus, eriala kood, üliõpilaste koguarv ning üliõpilaste arvu kogusumma protsentuaalselt.

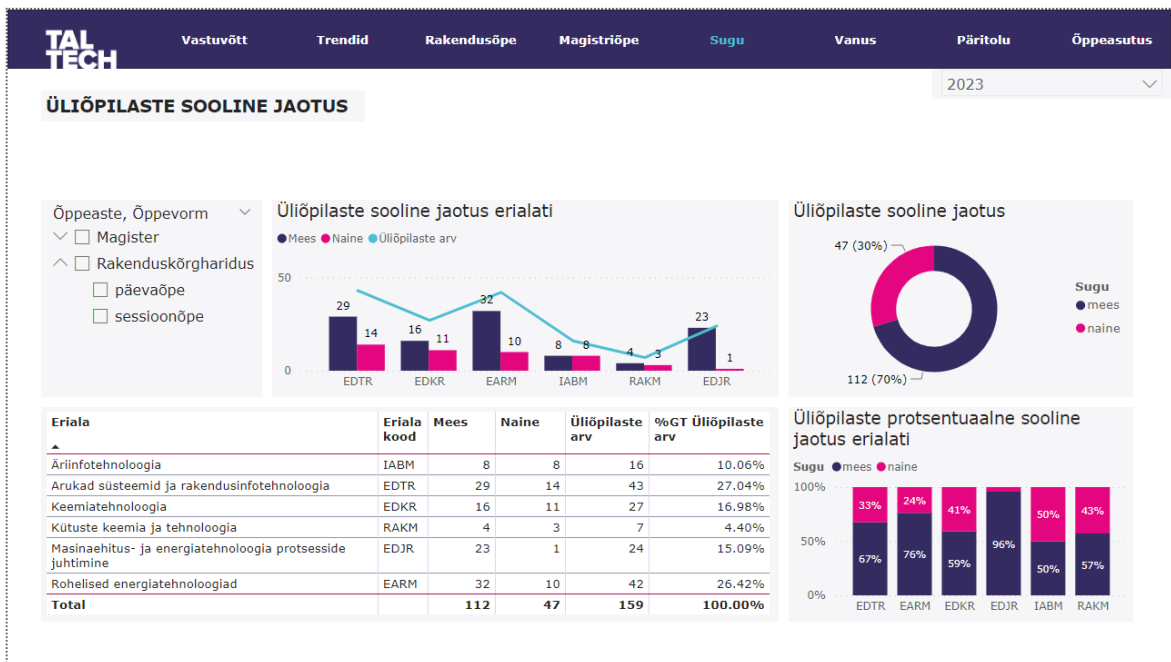


Joonis 4.16 Aruande lehekülg „Magistriõpe“

### 4.3.5 Üliõpilaste sooline jaotus

Aruande leheküljel „Sugu“ (vt Joonis 4.17), mis kannab pealkirja „Üliõpilaste sooline jaotus“, on esitatud Virumaa kolledži vastuvõtustatistika üliõpilaste soolise jaotuse järgi järgmiste visuaalide abil:

- **Ringdiagramm** „Üliõpilaste sooline jaotus“, mis visualiseerib üliõpilaste soolist jaotust numbriliselt ja protsentuaalselt.
- **Joon- ja virnastatud tulpdiagramm** „Üliõpilaste sooline jaotus erialati“, mis visualiseerib mees- ja naissoost üliõpilaste arvu erialade lõikes võrreldes üliõpilaste koguarvuga.
- **100% virnastatud tulpdiagramm** „Üliõpilaste protsentuaalne sooline jaotus erialati“, mis visualiseerib mees- ja naissoost üliõpilaste protsentuaalset jaotust erialade kaupa.
- **Tabel**, kus on toodud eriala nimetus, eriala kood, mees- ja naissoost üliõpilaste arv, üliõpilaste koguarv ning üliõpilaste arvu kogusumma protsentuaalselt.
- **Filtri visualisaal**, võimaldamaks kõiki lehekülje visuaale filtreerida vastavalt õppeastmele ja õppevormile.

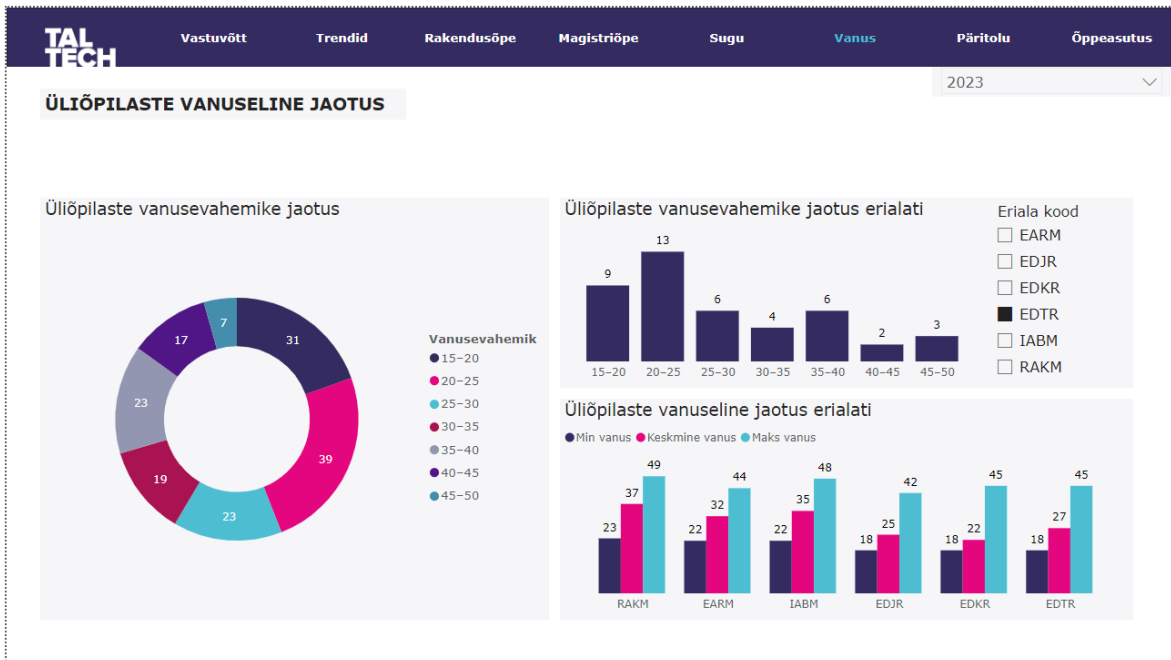


Joonis 4.17 Aruande lehekülg „Sugu“

### 4.3.6 Üliõpilaste vanuseline jaotus

Aruande leheküljel "Vanus" (vt Joonis 4.18), mis kannab pealkirja „Üliõpilaste vanuseline jaotus“, on esitatud TalTech Virumaa kolledži üliõpilaste vanuseline jaotus järgmiste visuaalide abil:

- **Ringdiagramm** „Üliõpilaste vanusevahemike jaotus“, mis visualiseerib üliõpilaste vanusevahemike jaotust numbriliselt ja protsentuaalselt.
- **Tulpdiagramm** „Üliõpilaste vanusevahemike jaotus erialati“, mille abil visualiseeritakse üliõpilaste arvu vanusevahemike kaupa. Selle kõrvale on loodud filtri visualisaal, võimaldamaks erialade kaupa filtreerimist.
- **Rühmitatud tulpdigramm** „Üliõpilaste vanuseline jaotus erialati“, mis kuvab üliõpilaste minimaalset, keskmist ja maksimaalset vanust erialade kaupa.



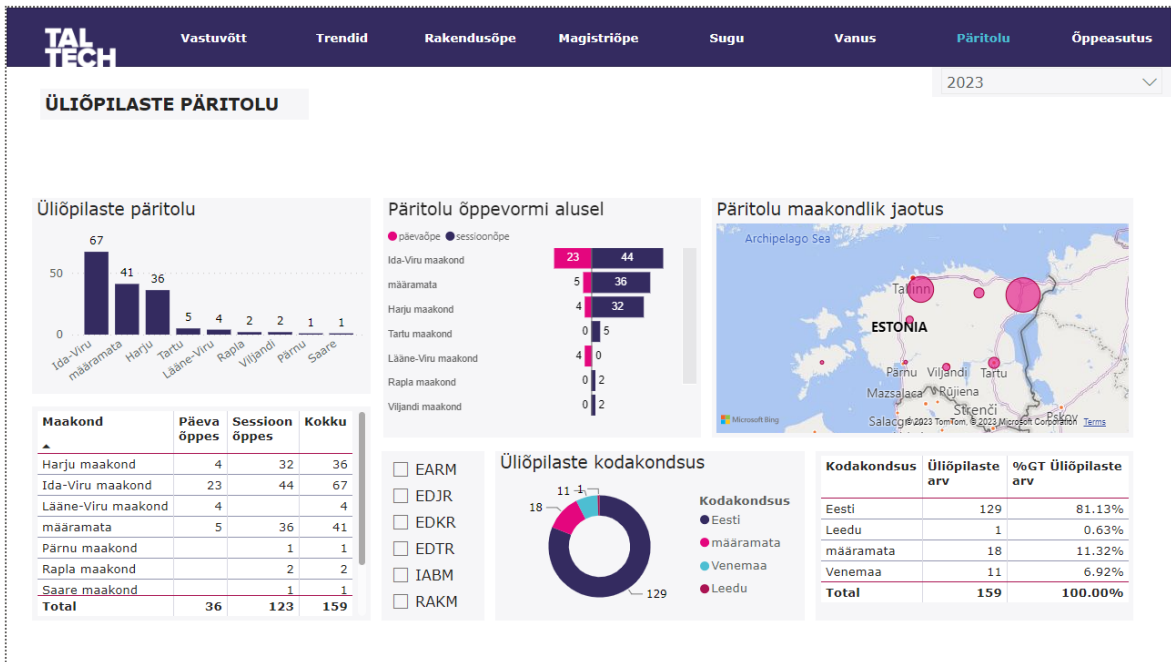
Joonis 4.18 Aruande lehekülg "Vanus"

#### 4.3.7 Üliõpilaste päritolu

Aruande leheküljel „Päritolu“ (vt Joonis 4.19) on esitatud TalTech Virumaa kolledži üliõpilaste kodakondsuse ja maakondliku päritolu statistika. Selleks on loodud järgmised seitse visuaali:

- **Tulpdiagramm** „Üliõpilaste päritolu“, mis kujutab üliõpilaste arvu maakondade kaupa.
- **Tornaadodiagramm** „Päritolu õppevormi alusel“, et visualiseerida päeva- ja sessioonõppe üliõpilaste arvu jagunemist maakondade vahel.
- **Geograafilise kaardi visuaal** „Päritolu maakondlik jaotus“, mis kuvab kaardil üliõpilaste maakondlikku jaotust erineva suurusega mullide abil.
- **Ringdiagramm** „Üliõpilaste kodakondsus“, mis visualiseerib üliõpilaste arvu jaotust kodakondsuse järgi.
- **Tabel** maakondadega, mis näitab maakonda, päeva- ja sessioonõppes olevate üliõpilaste arvu ning üliõpilaste koguarvu.
- **Tabel** kodakondsusega, mis esitab kodakondsuse, üliõpilaste arvu ja üliõpilaste koguarvu protsentuaalselt.
- **Filtrivisuaal**, mis võimaldab teisi visuaale filtreerida vastavalt eriala koodile.



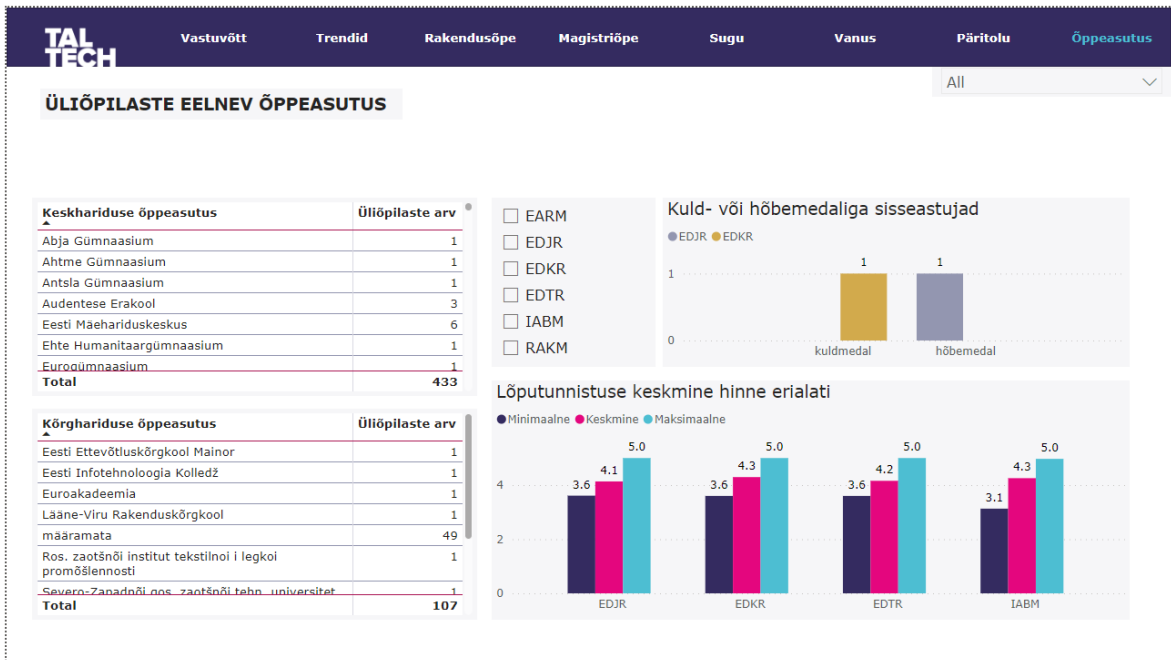


Joonis 4.19 Aruande lehekülg „Päritolu“

### 4.3.8 Üliõpilaste eelnev õppeasutus

Aruande leheküljel "Õppeasutus" (vt Joonis 4.20) on esitatud statistika TalTech Virumaa kolledžisse astunud üliõpilaste eelneva õppeasutuse kohta. Selleks loodi viis visuaali:

- **Tulpdiagramm** "Kuld- või hõbemedaliga sisseastujad", mis kujutab kuld- või hõbemedaliga sisseastunud üliõpilaste arvu erialade kaupa.
- **Rühmitatud tulpdiagramm** "Lõputunnistuse keskmine hinne erialati", mis visualiseerib erialade kaupa sisseastunud üliõpilaste minimaalset, keskmist ja maksimaalset keskmise hinde jaotust.
- **Tabel**, mis näitab keskhariduse õppeasutust ja üliõpilaste arvu.
- **Tabel**, mis näitab kõrghariduse õppeasutust ja üliõpilaste arvu.
- **Filtrivisuaal**, mis võimaldab leheküljel andmeid filtreerida eriala koodi alusel.



Joonis 4.20 Aruande lehekülj „Õppeasutus“

## 5. RAKENDUSE LOOMINE JA AUTOMATISEERIMINE

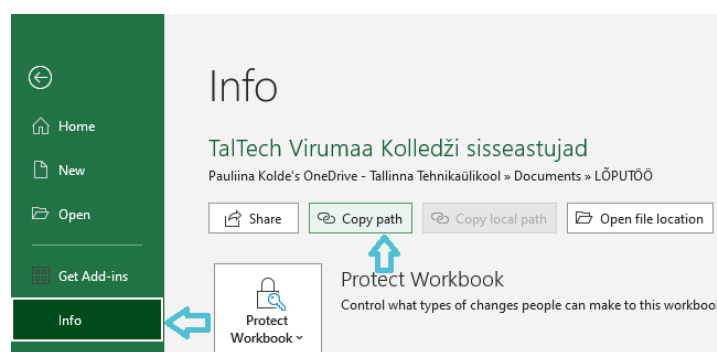
Power BI võimaldab Power BI Desktopi tarkvaras loodud aruandeid publitseerida <https://app.powerbi.com/> veebikeskkonnas, kus rakendusele saab jagada juurdepääsu teistele organisatsiooni liikmetele vaatamiseks. Ühtlasi luuakse rakenduse publitseerimisel rakenduse link, mis võimaldab rakenduse integreerida otse organisatsiooni enda veebileheküljele. Aruande rakendusena publitseerimiseks on tarvis see kõigepealt üles laadida töölauakeskkonda *Workspace*. [19]

Selleks et aruanne publitseerida rakendusena, mis uueneks automaatselt, tuli lähteandmed laadida üles OneDrive'i pilve ja muuta Power BI Desktopis aruande faili algallikaks OneDrive'i pilves asuv Exceli fail [20]. Kui aruanne sai rakendusena publitseeritud, siis oli tarvis sisse viia rakenduse automaatseks uuendamiseks vajalikud seadistused.

Käesolevas lõputöö peatükis kirjeldatakse rakenduse loomise ja automatiseerimise protsessi lähemalt ja antakse ülevaade rakenduse uuendamise protsessist.

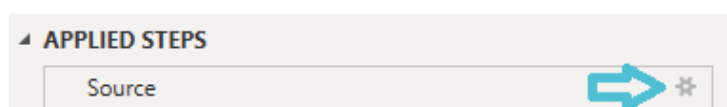
### 5.1 Faili algallika muutmine OneDrive'i asukohaks

Selleks et siduda Power BI Desktopis faili allikaks faili asukoht OneDrive'is, laaditi fail pilve ja avati seejärel Desktopi rakendusena, mis võimaldas kopeerida faili asukohta, valides menüüs *Info – Copy path* (vt Joonis 5.1).



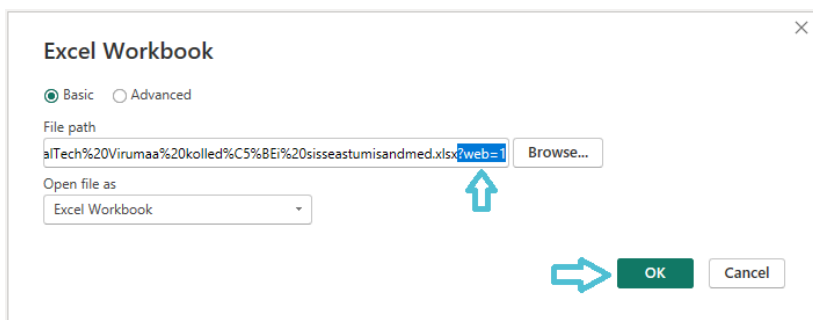
Joonis 5.1 OneDrive'i faili asukohta kopeerimine

Järgmise sammuna suunduti Power BI Desktopi, kus valides *Home* menüüst *Transform data*, suunduti edasi Power Query Editori redaktorisse. Power Query Editori rakendatud etappide jaotises *Applied Steps* tuli kõigi kolme tabeli puhul muuta lähteallikas *Source*, vajutades selle kõrval asuvale hammasratta ikoonile (vt Joonis 5.2).



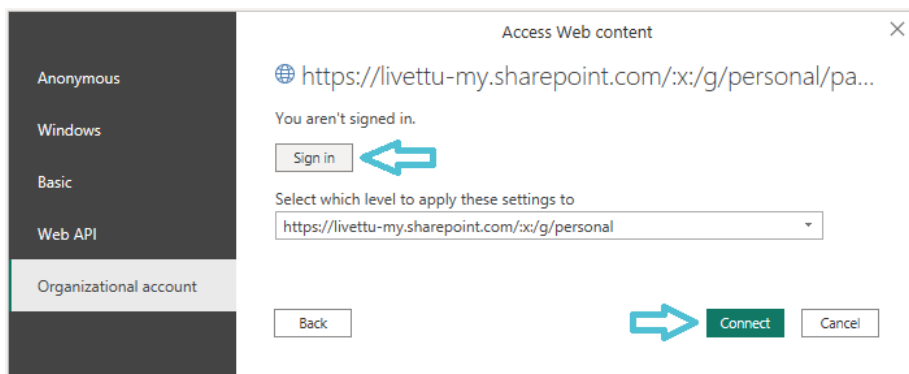
Joonis 5.2 Lähteallika muutmine

Avanendud hüpikaknas kopeeriti OneDrive'i faili asukoht *File path* lahtrisse, eemaldati lingilt lõpp „?web=1” ja vajutati OK (vt Joonis 5.3).



Joonis 5.3 Muudatused lähteallikas

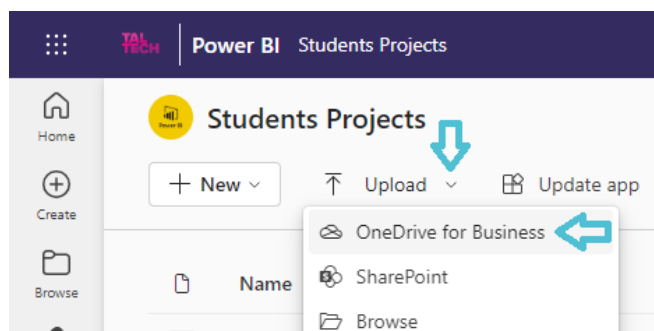
Selleks et Power BI saaks OneDrive pilvega ühenduse, tuli organisatsiooni kontoga sisse logida (vt Joonis 5.4).



Joonis 5.4 Organisatsiooni konto autentimine

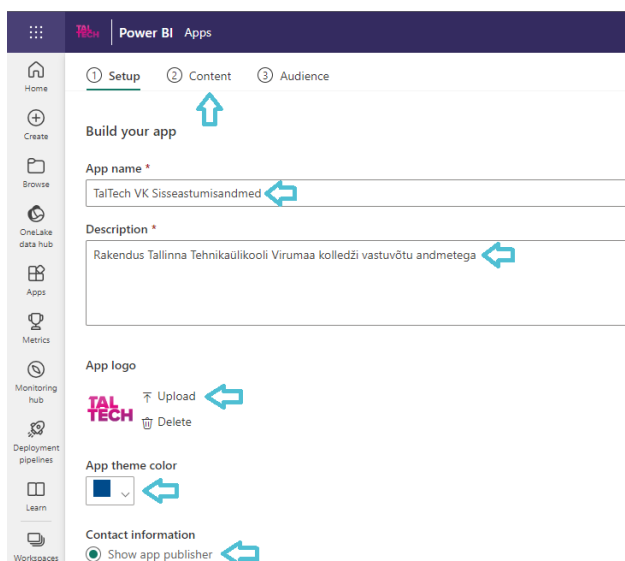
## 5.2 Aruande publitseerimine rakendusena

TalTech Virumaa kolledži üliõpilastele on Power BI veebikeskkonnas loodud töölaud Students Projects. Aruanne laaditi Power BI veebikeskkonda üles, valides külgmenüüst *Workspace* töölauda Student Projects ja seejärel töölauda menüüst *Upload – OneDrive for Business* (vt Joonis 5.5)



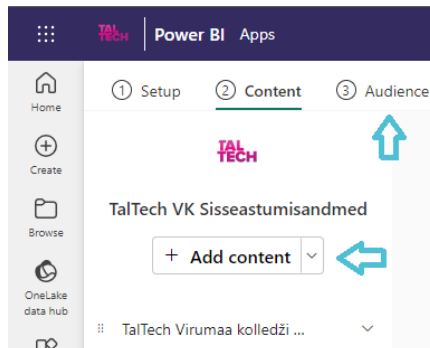
Joonis 5.5 Aruande laadimine OneDrive'ist Power BI veebikeskkonda

Seejärel uuendati töölauda rakendust, valides menüüst *Update App* ja määrates avanenud vaates *Setup* rakendusele nime, tutvustuse, logo pildi ja teemavärvi, suunduti edasi vaatesse *Content* (vt Joonis 5.6).



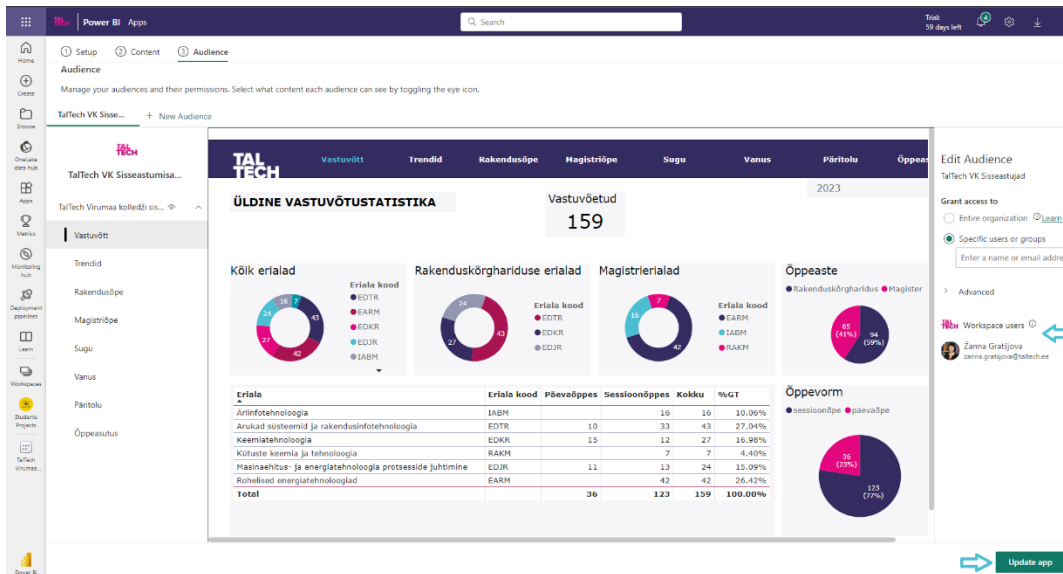
Joonis 5.6 Rakendusele andmete määramine

Vaates *Content* lisati rakendusele aruanne ja suunduti edasi vaatesse *Audience* (vt Joonis 5.7).



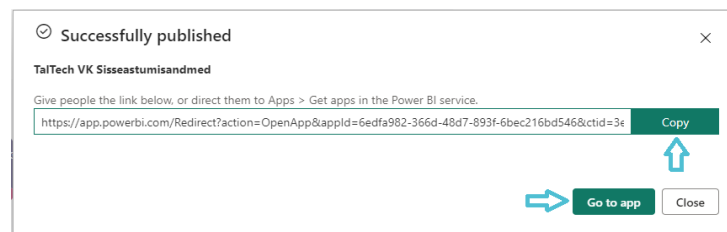
Joonis 5.7 Rakendusele aruande lisamine

Vaates *Audience* kontrolliti kõik aruande leheküljed üle, lisati vajalikud juurdepääsud, aktiveeriti rakenduse jagamise võimalus ja publitseeriti rakendus, vajutades nupule *Publish app* (vt Joonis 5.8).



Joonis 5.8 Rakenduse publitseerimine

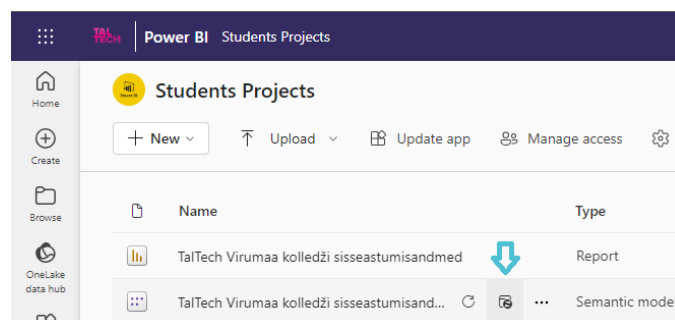
Rakenduse publitseerimise järel ilmus hüpinkaken rakenduse jagamise lingiga (vt Joonis 5.9).



Joonis 5.9 Rakenduse link

### 5.3 Rakenduse uuendamise automatiseerimine

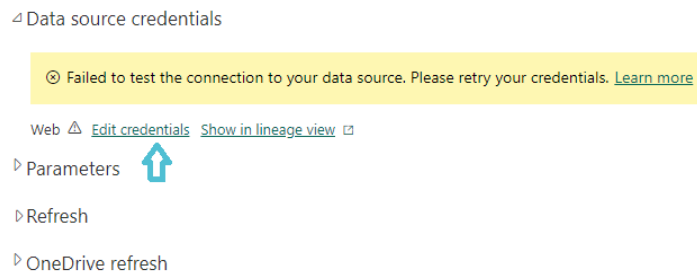
Laadides aruande Power BI veebikeskkonda, loodi aruandest automaatselt semantiline mudel, millele on võimalik seadetest lisada automaatne uuendamine, vajutades ikoonile *Schedule refresh* (vt Joonis 5.10).



Joonis 5.10 Semantilise mudeli uuendamine

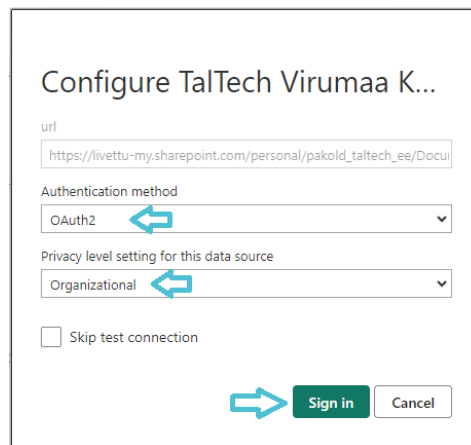
Selleks et automaatset uuendamist sätestada, tuleb kõigepealt *Data source credentials*

- *Edit credentials* kaudu andmeallikaga ühendus luua (vt Joonis 5.11).



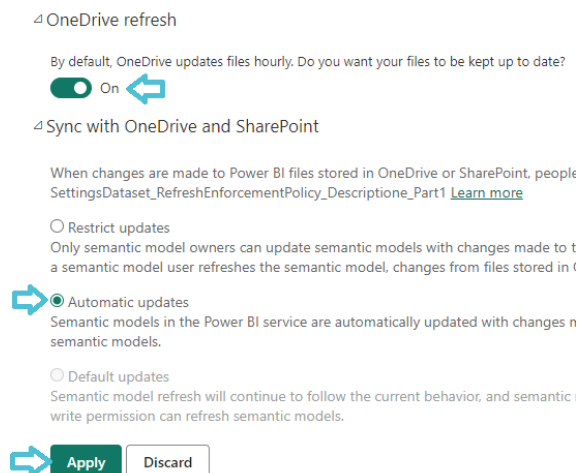
Joonis 5.11 Andmeallikaga autentimine

Organisatsiooni kontoga ühenduse loomiseks tuli autentimise meetodina kasutada *Oauth2* ja valida privaatsuse tasemeks *Organizational* (vt Joonis 5.12) [21].



Joonis 5.12 Autentimise meetodi määramine

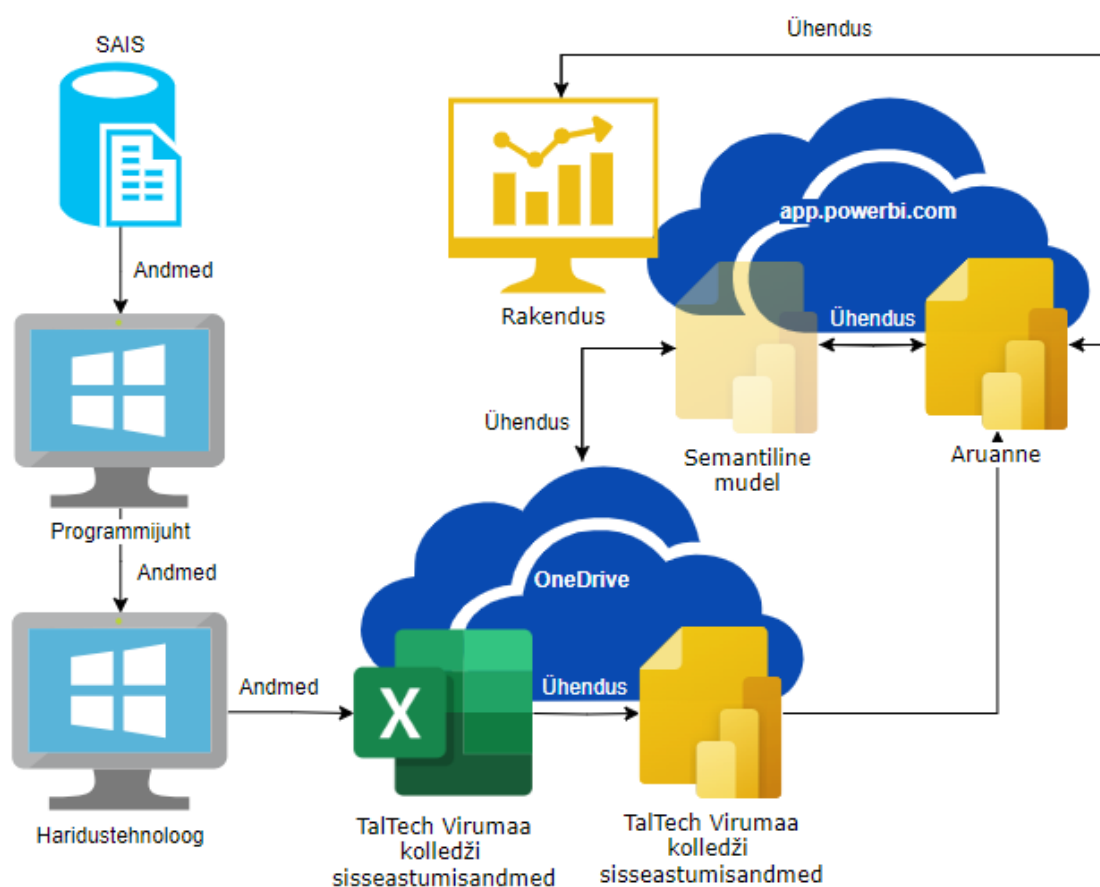
Seejärel lülitati sisse OneDrive'i automaatsed uuendused iga tunni järel ja rakendati semantilise mudeli sünkroniseerimine automaatseteks uuendusteks nupust *Apply* (vt Joonis 5.13).



Joonis 5.13 OneDrive'i sünkroniseerimine ja automaatsete uuenduste sisselülitamine

## 5.4 Rakenduse uuendamise protsessi kirjeldus

Rakenduse uuendamise protsess algab sellega, et programmijuhid saavad SAISist uued vastuvõtandmed. Seejärel edastatakse andmed haridustehnoloogile, kes lisab vastuvõtandmed OneDrive'is asuvasse Exceli faili nimega "TalTech Virumaa kolledži sisseastumisandmed". Kui Exceli fail on uuendatud, toimub umbes tunni aja jooksul automaatselt muudatus ka Power BI veebikeskkonnas, kus aruande semantilises failis ja aruandes "TalTech Virumaa kolledži sisseastumisandmed" tehakse uuendus vastavalt OneDrive's asuvale Exceli failile, mis rakendub siis ka Power BI rakendusele „TalTech VK sisseastumisandmed" (vt Joonis 5.14).



Joonis 5.14 Rakenduse uuendamise protsessi vooskeem



## KOKKUVÕTE

Käesoleva lõputöö peamiseks eesmärgiks oli automatiseerida TalTech Virumaa kolledži vastuvõtuandmete visualiseerimise protsess BI tarkvara abil, et hõlbustada kolledži vastuvõtustatistika analüüsimist. Power BI tarkvara valiti lähtudes tema turuliidri positsioonist ja Tallinna Tehnikaülikooli töötajate ligipääsetavusest tarkvarale.

Lõputöö autori ülesanneteks kujunes TalTech Virumaa kolledži probleemi kaardistamine, TalTech Virumaa kolledži aastate 2019–2023 vastuvõtuandmete visualiseerimiseks ette valmistamine, andmete visualiseerimine, aruande kujundamine ja aruande publitseerimine automaatselt uuendatava rakendusena.

Käesoleva lõputöö esimeses peatükis andis autor ülevaate lõputöö teema aktuaalsusest, kirjeldas lähemalt TalTech Virumaa kolledži vajadust vastuvõtuandmete visualiseerimise protsessi automatiseerimiseks ja põhjendas Power BI tarkvara valikut.

Andmete ettevalmistamise protsess hõlmas esialgse andmestikuga tutvumist MS Excelis, Power BI Desktopi Power Query Editori redaktoris põhjalikku andmestiku korrastamist ja andmemudeli loomist ning Power BI Desktopis andmete visualiseerimiseks oluliste arvutatud mõõdikute ja veergude lisamist.

Sisseastumisandmed visualiseeriti ja koostati Power BI aruanne, mis on teemade kaupa jaotatud kaheksaks leheküljeks: „Vastuvõtt“, „Trendid“, „Rakendusõpe“, „Magistriõpe“, „Sugu“, „Vanus“, „Päritolu“ ja „Õppeasutus“. Kokku kasutati aruandes andmete visualiseerimiseks kaheksat ringdiagrammi, kolme sektordiagrammi, nelja kaardi visuaali, kolme tulpdigrammi, kolme rühmitatud tulpdigrammi, kuute joon- ja virnastatud tulpdigrammi, ühte 100% virnastatud tulpdigrammi, ühte tornaadodiagrammi, ühte geograafilise kaardi visuaali ja kaheksat tabelit.

Lõpetuseks publitseeriti aruanne Power BI veebikeskkonnas rakendusena, automatiseeriti rakenduse uuendamine ja koostati rakenduse uuendamise protsessi kirjeldus.

Käesolevas lõputöös on autor dokumenteerinud andmete ettevalmistamise, visuaalide loomise, aruande koostamise ja rakenduse uuendamise protsessi piisava detailsusega, võimaldamaks tulevikus TalTech Virumaa kolledži vastuvõtu statistika eest vastutaval isikul probleemideta integreerida uusi andmeid aruandesse ja vajadusel muuta aruandes tehtud visualiseerimisi.

Autor usub, et lõputöö käigus loodud vastuvõtuandmete visualiseerimise automatiseeritud lahendus mitte ainult ei võimalda tulevikus aja kokkuhoidu, vaid pakub ka mitmeid praktilisi eeliseid, parandades vastuvõtuandmete kättesaadavust ja haldamist. Loodud rakendus koondab kõik vastuvõtuandmed ühte kohta, esitades need

standardsel kujul. Samuti võimaldab see andmeid filtreerida, muutes olulise info kiiresti ja lihtsasti leitavaks.

## SUMMARY

As society experiences a surge in data volume, proficiency in data analysis has become predominant. An integral part of data analysis is data visualization, which facilitates understanding of the data. As technology progresses, so do the tools at our disposal for efficient data analysis. At TalTech Virumaa college, admission data has been visualized each academic year using MS Office software, which is not time efficient and is becoming outdated practice.

The main objective of Pauliina Kolde's thesis, "Visualization of TalTech Virumaa College Admission Data Using Power BI Software," was to automate the visualization of admission data at TalTech Virumaa college using MS Power BI software to facilitate the annual process of visualizing admission data at the college. Power BI software was chosen for automating data processing, considering its market leading position and accessibility to Tallinn University of Technology employees.

The tasks undertaken in the thesis included organizing admission data for the years 2019–2023 using MS Excel. To automate the data cleaning and organization process, the dataset was processed into a data model using Power Query Editor in Power BI Desktop, with calculated measures and columns added using DAX language.

Admission data was visualized using appropriate visuals and organized into an eight page Power BI report divided by topics: "Admission," "Trends," "Applied Studies," "Master's Studies," "Gender," "Age," "Origin," and "Educational Institution." In total, the report used: eight pie charts, three donut charts, four card visuals, six clustered column charts, six line and stacked column charts, one 100% stacked column chart, one tornado diagram, one geographical map visual, and eight tables for data visualization.

Finally, the report was published on the Power BI web environment as an application with automated updates and a description of the application update process was compiled.

Throughout the thesis, the author documented the data preparation, creation process of all visuals used in the report, report design and structure, and application update processes in sufficient detail to ensure that the individual responsible for TalTech Virumaa college's admission statistics will be able to seamlessly integrate new data and make necessary visual adjustments.

The author believes that the automated solution developed during this thesis not only saves time in the future but also provides practical advantages, enhancing accessibility and improving the management of admission data. The created application consolidates admission data for all years into one place in a standardized format, providing the added functionality of data filtering for swift access to essential information.

## KASUTATUD KIRJANDUSE LOETELU

1. Masso, A., Tiidenberg, K., Siibak, A. Kuidas mõista andmestunud maailma? Metodoloogiline teejuht, Tallinn: Tallinna Ülikooli kirjastus, 2020.
2. Tiit, E.-M. STATISTICS IN THE PAST, THE PRESENT AND THE FUTURE, Eesti Statistika Kvartalikiri, nr 2, lk 17-22, 2016.
3. Bharadiya, J. P. A Comparative Study of Business Intelligence and Artificial Intelligence with Big Data Analytics, American Journal of Artificial Intelligence, 2023, nr 7(1), lk 24-30.
4. Schlegel, K., Sun, J., Pidsley, D., Ganeshan, A., Fei, F., Popa, A. Miclaus, R., Macari, E., Quinn, K., Long, C. Magic Quadrant for Analytics and Business Intelligence Platforms. [Online] [https://www.gartner.com/doc/reprints?id=1-2955ETOT&ct=220215&st=sb?ocid=lp\\_pg398450\\_gdc\\_comm\\_az](https://www.gartner.com/doc/reprints?id=1-2955ETOT&ct=220215&st=sb?ocid=lp_pg398450_gdc_comm_az) (20.12.2023).
5. Edmond, S., Crabtree, M. Power BI vs Tableau: Which is The Better Business Intelligence Tool in 2024? [Online] <https://www.datacamp.com/blog/power-bi-vs-tableau-which-one-should-you-choose> (20.12.2023).
6. K. Haan, The Best Data Analytics Tools Of 2023. [Online] <https://www.forbes.com/advisor/business/software/best-data-analytics-tools> (20.12.2023).
7. How to Clean Data in Power BI. [Online] <https://zebrabi.com/guide/how-to-clean-data-in-power-bi> (20.12.2023).
8. The Complete Guide to Power Query. [Online] <https://www.howtoexcel.org/the-complete-guide-to-power-query> (20.12.2023).
9. What is Power Query? [Online] Available: <https://powerquery.microsoft.com/en-us> (20.12.2023).
10. Power BI - Star schema or single table. [Online] <https://www.sqlbi.com/articles/power-bi-star-schema-or-single-table> (20.12.2023).
11. Bothma, J. Data Modeling in Power BI Tutorial. [Online] <https://www.datacamp.com/tutorial/data-modeling-in-power-bi-tutorial> (20.12.2023).
12. Understand star schema and the importance for Power BI. [Online] <https://learn.microsoft.com/en-us/power-bi/guidance/star-schema> (20.12.2023).

13. Create and manage relationships in Power BI Desktop. [Online] <https://learn.microsoft.com/en-us/power-bi/transform-model/desktop-create-and-manage-relationships> (20.12.2023).
14. Model relationships in Power BI Desktop. [Online] <https://learn.microsoft.com/en-us/power-bi/transform-model/desktop-relationships-understand> (20.12.2023).
15. Granlund, S. Mis on Power BI DAX ja kuidas seda andmetega töötamisel kasutada? [Online] <https://www.finantsuudised.ee/uudised/2022/03/08/mis-on-power-bi-dax-ja-kuidas-seda-andmetega-tootamisel-kasutada-> (20.12.2023).
16. Malamed, C. Visual Design Solutions: Principles and Creative Inspiration for Learning Professionals, Hoboken, New Jersey: John Wiley & Sons, Inc, 2015.
17. TALLINNA TEHNIKAÜLIKOOLI STIILIRAAMAT. [Online] [https://haldus.taltech.ee/sites/default/files/2019-09/TalTech\\_CVI.pdf](https://haldus.taltech.ee/sites/default/files/2019-09/TalTech_CVI.pdf) (20.12.2023).
18. Gradient Generator. [Online] <https://colordesigner.io/gradient-generator> (20.12.2023).
19. Publish an app in Power BI. [Online] <https://learn.microsoft.com/en-us/power-bi/collaborate-share/service-create-distribute-apps> (20.12.2023).
20. Refresh a semantic model created from an Excel workbook on OneDrive or SharePoint Online. [Online] <https://learn.microsoft.com/en-us/power-bi/connect-data/refresh-excel-file-onedrive> (20.12.2023).
21. Use OneDrive for work or school links in Power BI Desktop. [Online] <https://learn.microsoft.com/en-us/power-bi/connect-data/desktop-use-onedrive-business-links> (20.12.2023).