

Univerzitet Singidunum

Fakultet za informatiku i računarstvo

Obrada podataka o prodaji kafe

Projekat za prvi kolokvijum iz predmeta Veliki skupovi podataka

Profesor: Dalibor Radovanović

Predmet: Veliki skupovi podataka

Student: Nikola Nenadović

Broj indeksa: 2023205687

1 Uvod

1.1 QlikView

QlikView je softverski alat iz oblasti analitike podataka i poslovne inteligencije (Business Intelligence) namenjen analizi i vizuelnoj reprezentaciji podataka. QlikView je razvijen od strane kompanije Qlik (prethodno je nosila ime QlikTech), osnovane 1993. godine od strane Björn Berga i Staffana Gestrelusa u švedskom gradu Lund. Qlik je razvio softverske alate za replikaciju podataka (Qlik Replicate), za organizaciju podataka (Qlik Catalog), QlikCompose program za automatizaciju skladišta podataka, i mnoge druge. Omogućava nam da formiramo tabele, kreiramo izveštaje i različite vrste grafikona i dijagrama namenjenih za predstavljanje podataka iz više izvora. Pomoću QlikView-a možemo vršiti analitiku podataka sa društvenih mreža (npr, možemo generisati detaljne vizuelne grafike i analitičke izveštaje na osnovu raznovrsnih podataka „povučenih“ (scrapeovanih) sa društvenih mreža), analizirati prihode i druge mikroekonomske pokazatelje iz različitih datasetova o prodaji proizvoda, vršiti predviđanja i prognoze koje mogu biti od velike pomoći pri donošenju poslovnih odluka, kao i mnoge druge funkcije iz oblasti analize podataka. QlikView omogućava veliku fleksibilnost u kreiranju izveštaja, širok izbor različitih postupaka analitike uz podršku velikog broja funkcija.

U nedostatke QlikView softvera možemo ubrojiti, ponekad usporen proces učitavanja podataka ukoliko je u pitanju dataset sa velikom količinom podataka, dok obrada datasetova sa velikim brojem atributa i vrednosti može dovesti do usporenog rada, pa čak i prekida programa.

2 Opis dataseta

U projektu smo analizirali i predstavili podatke iz *Coffee Shop Sales* dataseta. Dataset sadrži podatke o prodaji kafe i drugih srodnih proizvoda, koji su sakupljeni iz tri kafeterije sa lokacijom u Njujorku, u Sjedinjenim Američkim Državama. Dataset je postavljen na sajtu Kaggle.com, a izvor podataka je Maven Analytics servis. U datasetu su predstavljeni podaci o prodaji od januara do juna 2023. godine. Dataset sadrži dve tabele, *Transactions* i *Customers* i sastoji se iz ukupno 21 kolone (atributa ili featurea), na osnovu kojih se mogu jasno utvrditi obrasci ponašanja kupaca. Kolone u *Transactions* tabeli su:

- `transaction_id` : Identifikacioni broj koji jedinstveno određuje svaku transakciju
- `transaction_date` : Datum svake transakcije u format MM/DD/YY
- `transaction_time` : Tačno vreme svake kupovine predstavljeno u formatu HH:MM:SS
- `transaction_qty` : Prodata količina jednog artikla predstavljena numeričkim tipom podataka
- `store_id` : Identifikacioni broj fizičke prodavnice (kafeterije)
- `store_location` : Lokacija kafeterije u kojoj je obavljena kupovina
- `product_id` : Jedinstveni identifikacioni broj svakog proizvoda
- `unit_price` : Cena proizvoda predstavljena u numeričkom formatu
- `product_category` : Kategorija proizvoda predstavljena tekstualnim tipom podataka

- `product_type` : Ime prodatog proizvoda
- `product_detail`: Detaljniji opis prodatog proizvoda
- `Revenue`: Prihod po svakoj transakciji
- `Month`:Mesec u kojem je izvršena kupovina predstavljen brojem iz intervala [1,12]
- `Weekday`: Dan u nedelji u kojem je izvršena transakcija predstavljen brojem iz intervala [1,7]
- `Hour`: Sat u kojem je izvršena kupovina predstavljena bročanom vrednošću iz intervala [6, 20]
- `Month1`: Mesec u kojem je obrađena transakcija predstavljen tekstualnim tipom podataka(Jan, Feb, itd.)
- `Weekday1`: Dan u nedelji u kojem je izvršena transakcija predstavljen tekstualnim tipom podataka(Mon, Tue, Wed, itd.)

Customers tabela dataseta Coffee Shop Sales sadrži sledeće kolone:

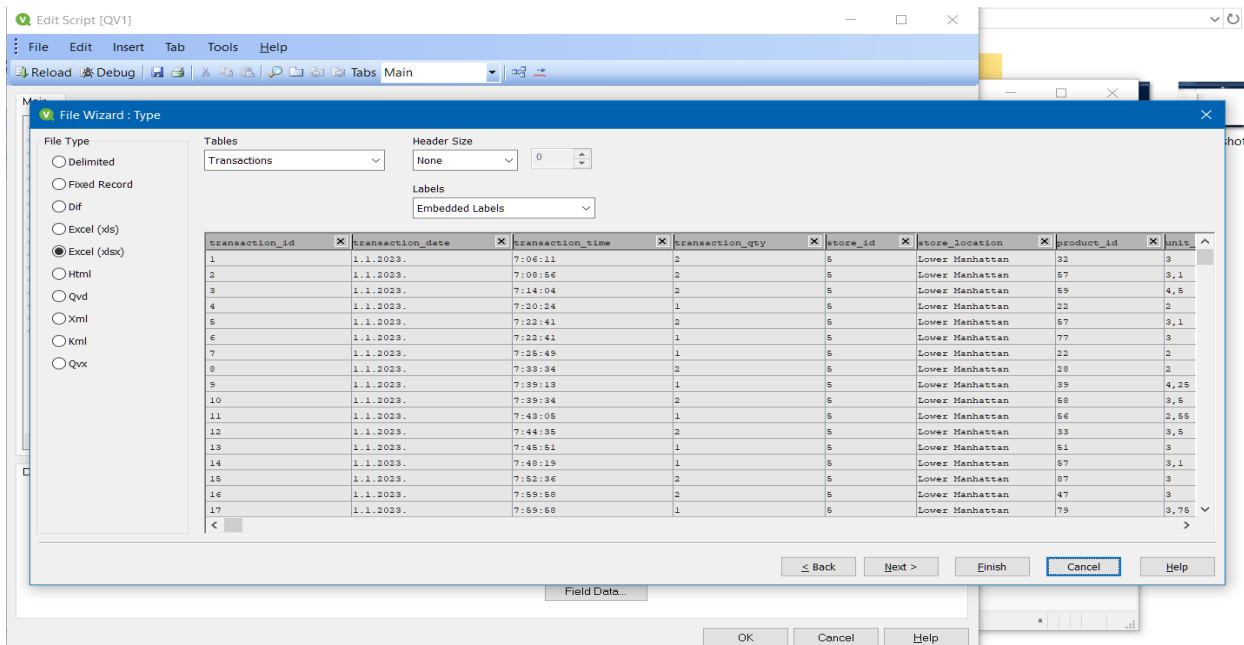
- `purchase_id`:Numerička vrednost koja jedinstveno određuje svaku kupovinu
- `customer_name`:Ime i prezime kupca
- `payment_method`:Način plaćanja(gotovina, kartica ili mobilni telefon)
- `invoice_number`: Jedinstveni broj računa

Sam dataset je kvalitetno urađen,sa malim brojem NULL vrednosti, sadrži različite tipove podataka a same vrednosti su raznovrsne što omogućava kvalitetnu analizu podataka i pouzdane rezultate na osnovu kojih možemo vršiti dalje predikcije i povlačiti poslovne poteze.Na osnovu podataka iz dataseta možemo prikazati koji proizvodi su najpopularniji kod kupaca, koje kategorije proizvoda nam donose najviše prihoda, kada kupci obavljaju najveći broj kupovina, odnosno u kom vremenskom intervalu ostvarujemo najveći broj transakcija,koje prodavnice ostvaruju najviše prihoda, itd.

3 Ubacivanje i čišćenje podataka u QlikView softveru

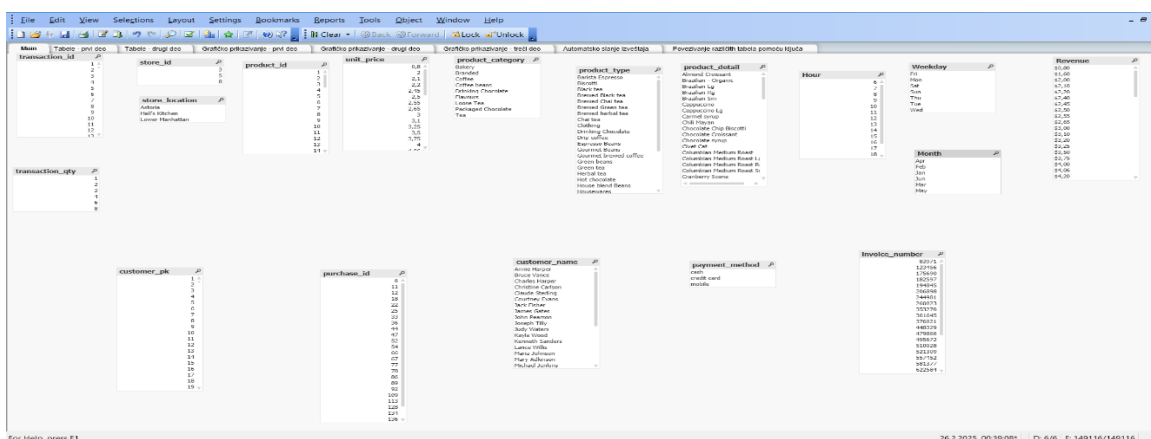
Početak izrade projekta sastoji se od otvaranja QlikView programa i učitavanja dataseta.Nakon pokretanja QlikViewa, u gornjem meniju biramo funkciju *Edit Script* i tako pokrećemo proces prepoznavanja dataseta od strane QlikView softvera.Klikom na dugme *Table files* otvara nam se prozor u kojem je potrebno pronaći dataset na našem sistemu.Nakon učitavanja dataseta otvara se novi prozor u kojem vršimo izbor kolona ill atributa, čije vrednosti podataka će biti analizirane i prezentovane na projektu. U prvom koraku učitvanja dataseta može se vršiti izbor formata fajla dataseta(CSV, xls., xlsx,), dodeljivane pseudonima kolonama dataseta, pozicioniranje imena kolona, itd..Za potrebe projekta eliminisaćemo četiri kolone dataseta:*transaction_time*, *transaction_date*, *Month* i *Weekday*. Kolonama *Month1* i *Weekday1* dodeljuju se pseudonimi(aliasi) *Month* i *Weekday* radi bolje preglednosti podataka.Ostale procedure u

početnom učitavanju dataseta su rutinske prirode, iz razloga što je dataset u odličnom stanju za obradu i analizu podataka.



Slika 1 – Izbor odgovarajućih tabela, kolona, oznaka, itd.

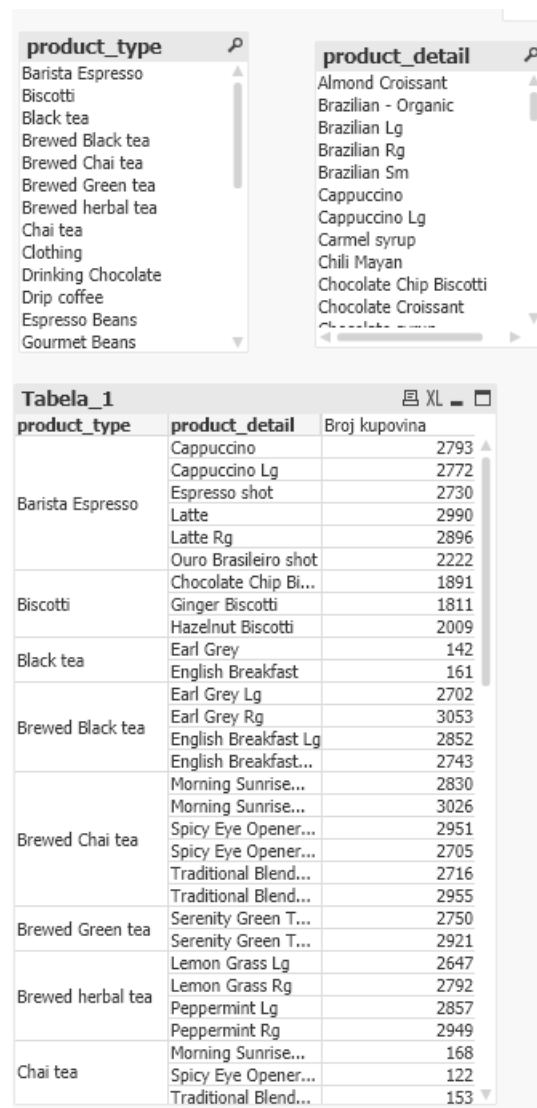
Nakon čišćenja i transformacije podataka dataseta, klikom na dugme Reload vrši se prenos podataka u QlikView softver i biranje odgovarajućih kolona dataseta. Zatim, otvara se početni tab našeg projekta sa imenom Main, gde se nalaze svi atributi podataka koje smo odabrali za analizu i predstavljanje.



Slika 2 – Main ili početni prozor našeg projekta

4 Kreiranje kros i pivot tabele u Qlikview softveru

Za potrebe projekta kreirano je ukupno šest tabela, jedna tabela iz kategorije Straight i pet tabela iz kategorije Pivot. U prvoj Straight tabeli, pod imenom Tabela_1, dimenzije tabele su atributi (kolone) product_type i product_detail, koji predstavljaju tip proizvoda i detalje pojedinačnog proizvoda koje se prodaje. Kao izraz (Expression) za tabelu, izračunao sam koliko je puta svaki pojedinačni proizvod bio predmet transakcije.



The screenshot shows the QlikView interface. At the top, there are two dimension lists: 'product_type' and 'product_detail'. Below them is a table view titled 'Tabela_1' with three columns: 'product_type', 'product_detail', and 'Broj kupovina'. The table displays data for various product types and their details, with the number of purchases for each.

| product_type | product_detail | Broj kupovina |
|-------------------|----------------------|---------------|
| Barista Espresso | Cappuccino | 2793 |
| | Cappuccino Lg | 2772 |
| | Espresso shot | 2730 |
| | Latte | 2990 |
| | Latte Rg | 2896 |
| | Ouro Brasileiro shot | 2222 |
| Biscotti | Chocolate Chip Bi... | 1891 |
| | Ginger Biscotti | 1811 |
| | Hazelnut Biscotti | 2009 |
| Black tea | Earl Grey | 142 |
| | English Breakfast | 161 |
| Brewed Black tea | Earl Grey Lg | 2702 |
| | Earl Grey Rg | 3053 |
| | English Breakfast Lg | 2852 |
| | English Breakfast... | 2743 |
| Brewed Chai tea | Morning Sunrise... | 2830 |
| | Morning Sunrise... | 3026 |
| | Spicy Eye Opener... | 2951 |
| | Spicy Eye Opener... | 2705 |
| Brewed Green tea | Traditional Blend... | 2716 |
| | Traditional Blend... | 2955 |
| | Serenity Green T... | 2750 |
| Brewed herbal tea | Serenity Green T... | 2921 |
| | Lemon Grass Lg | 2647 |
| | Lemon Grass Rg | 2792 |
| | Peppermint Lg | 2857 |
| Chai tea | Peppermint Rg | 2949 |
| | Morning Sunrise... | 168 |
| | Spicy Eye Opener... | 122 |
| | Traditional Blend... | 153 |

Slika 3 - Tabela_1 prikazana zajedno sa njenim dimenzijama

U drugoj tabeli iz kategorije Pivot, koja nosi ime Tabela_2, dimenzije tabele su atributi product_category i product_type, koji tekstualnim tipom podataka opisuju kategoriju proizvoda i

bliže opisuju pojedinačni proizvod. Za izraze (Expressions) tabele, prikazani su broj transakcija, količina prodatih proizvoda i prihod kafeterija sumiran po tipu proizvoda.

product_category

- Bakery
- Branded
- Coffee
- Coffee beans
- Drinking Chocolate
- Flavours
- Loose Tea
- Packaged Chocolate
- Tea

product_type

- Barista Espresso
- Biscotti
- Black tea
- Brewed Black tea
- Brewed Chai tea
- Brewed Green tea
- Brewed herbal tea
- Chai tea
- Clothing
- Drinking Chocolate
- Drip coffee
- Espresso Beans
- Gourmet Beans

product_detail

- Almond Croissant
- Brazilian - Organic
- Brazilian Lg
- Brazilian Rg
- Brazilian Sm
- Cappuccino
- Cappuccino Lg
- Caramel syrup
- Chili Mayan
- Chocolate Chip Biscotti
- Chocolate Croissant
- Chocolate...

Tabela_2

product_category

product_type

Broj transakcija

Količina prodatih proizvoda

Prihod

| | | | | |
|--------------------|---------------------|-------|-------|-----------|
| Bakery | Biscotti | 5711 | 5788 | 19.794 \$ |
| | Pastry | 6912 | 6961 | 25.656 \$ |
| | Scone | 10173 | 10465 | 36.866 \$ |
| Branded | Clothing | 221 | 221 | 6.163 \$ |
| | Housewares | 526 | 555 | 7.444 \$ |
| Coffee | Barista Espresso | 16403 | 24943 | 91.406 \$ |
| | Drip coffee | 8477 | 12891 | 31.984 \$ |
| | Gourmet brewed... | 16912 | 25973 | 70.035 \$ |
| | Organic brewed C... | 8489 | 13012 | 37.747 \$ |
| | Premium brewed... | 8135 | 12431 | 38.781 \$ |
| Coffee beans | Espresso Beans | 319 | 319 | 5.560 \$ |
| | Gourmet Beans | 366 | 366 | 6.798 \$ |
| | Green beans | 134 | 134 | 1.340 \$ |
| | House blend Beans | 183 | 183 | 3.294 \$ |
| | Organic Beans | 415 | 420 | 8.510 \$ |
| | Premium Beans | 336 | 406 | 14.584 \$ |
| Drinking Chocolate | Hot chocolate | 11468 | 17457 | 72.416 \$ |
| Flavours | Regular syrup | 4979 | 7606 | 6.085 \$ |
| | Sugar free syrup | 1811 | 2905 | 2.324 \$ |
| Loose Tea | Black tea | 303 | 303 | 2.712 \$ |
| | Chai tea | 443 | 443 | 4.301 \$ |
| | Green tea | 159 | 159 | 1.471 \$ |
| | Herbal tea | 305 | 305 | 2.730 \$ |
| Packaged Chocolate | Drinking Chocolate | 266 | 266 | 2.728 \$ |
| | Organic Chocolate | 221 | 221 | 1.680 \$ |
| Tea | Brewed Black tea | 11350 | 17462 | 47.932 \$ |
| | Brewed Chai tea | 17183 | 26250 | 77.082 \$ |
| | Brewed Green tea | 5671 | 8697 | 23.853 \$ |
| | Brewed herbal tea | 11245 | 17328 | 47.540 \$ |

Slika 4 - Tabela_2

Dimenzije treće Pivot tabele (Tabela_3) su atributi Month i Weekday, koji tekstualnim tipom podataka predstavljaju mesece u godini i dane u nedelji. Kao izrazi tabele izračunati su broj transakcija po danu, količina prodatih proizvoda po danu i dnevni prihod.

Month
Apr
Feb
Jan
Jun
Mar
May

unit_price
0,8
2
2,1
2,2
2,45
2,5
2,55
2,65

Weekday
Fri
Mon
Sat
Sun
Thu
Tue
Wed

| Month | Weekday | Broj transakcija | Količina prodatih... | Dnevni prihod |
|-------|---------|------------------|----------------------|---------------|
| Apr | Fri | 3.201 | 4.552 | 15.193 \$ |
| | Mon | 3.415 | 4.907 | 16.423 \$ |
| | Sat | 4.116 | 5.962 | 19.310 \$ |
| | Sun | 4.279 | 6.073 | 20.039 \$ |
| | Thu | 3.417 | 5.001 | 15.717 \$ |
| | Tue | 3.390 | 4.933 | 15.789 \$ |
| | Wed | 3.517 | 5.041 | 16.471 \$ |
| Feb | Fri | 2.358 | 3.420 | 10.735 \$ |
| | Mon | 2.373 | 3.484 | 11.093 \$ |
| | Sat | 2.321 | 3.333 | 10.767 \$ |
| | Sun | 2.433 | 3.473 | 11.375 \$ |
| | Thu | 2.325 | 3.281 | 10.887 \$ |
| | Tue | 2.199 | 3.130 | 10.203 \$ |
| | Wed | 2.350 | 3.429 | 11.085 \$ |
| Jan | Fri | 2.291 | 3.339 | 10.653 \$ |
| | Mon | 2.814 | 3.965 | 13.239 \$ |
| | Sat | 2.206 | 3.120 | 10.422 \$ |
| | Sun | 2.725 | 3.955 | 12.743 \$ |
| | Thu | 2.283 | 3.282 | 10.523 \$ |
| | Tue | 2.766 | 3.968 | 13.740 \$ |
| | Wed | 2.229 | 3.241 | 10.358 \$ |
| Jun | Fri | 5.960 | 8.533 | 28.198 \$ |
| | Mon | 4.810 | 6.946 | 22.520 \$ |
| | Sat | 4.683 | 6.760 | 22.817 \$ |
| | Sun | 4.786 | 6.904 | 22.185 \$ |
| | Thu | 5.812 | 8.391 | 27.251 \$ |
| | Tue | 4.801 | 7.038 | 22.259 \$ |
| | Wed | 4.500 | 6.370 | 21.257 \$ |
| Mar | Fri | 3.486 | 5.000 | 16.253 \$ |
| | Mon | 2.816 | 4.111 | 13.182 \$ |
| | Sat | 2.732 | 3.950 | 12.783 \$ |
| | Sun | 2.738 | 3.925 | 12.684 \$ |
| | Thu | 3.486 | 4.940 | 16.136 \$ |
| | Tue | 2.596 | 3.647 | 12.119 \$ |
| | Wed | 3.375 | 4.833 | 15.678 \$ |
| May | Fri | 4.405 | 6.363 | 20.341 \$ |
| | Mon | 5.415 | 7.818 | 25.221 \$ |
| | Sat | 4.452 | 6.489 | 20.795 \$ |
| | Sun | 4.135 | 5.852 | 19.305 \$ |
| | Thu | 4.331 | 6.267 | 20.254 \$ |

Slika 5 – Tabela_3

U četvrtoj Pivot tabeli, kao dimenzije tabele prikazani su atributi Weekday i Hour. Kao izraz tabele, izračunat je i prikazan broj transakcija u svakom satu rada kafeterije, kao i prihod u dolarima sumiran po satima radnog vremena kafeterije.

Weekday

Fri

Mon

Sat

Sun

Thu

Tue

Wed

Hour

6

7

8

9

10

11

12

13

14

15

16

17

Tabela_4

| Weekday | Hour | Broj transakcija | Prihod |
|---------|------|------------------|-----------|
| Fri | 6 | 707 | 3.386 \$ |
| | 7 | 1989 | 9.386 \$ |
| | 8 | 2544 | 11.747 \$ |
| | 9 | 2707 | 12.750 \$ |
| | 10 | 2761 | 13.156 \$ |
| | 11 | 1405 | 6.828 \$ |
| | 12 | 1231 | 5.685 \$ |
| | 13 | 1199 | 5.499 \$ |
| | 14 | 1265 | 5.840 \$ |
| | 15 | 1295 | 6.098 \$ |
| | 16 | 1257 | 5.655 \$ |
| | 17 | 1226 | 5.687 \$ |
| | 18 | 1149 | 5.140 \$ |
| | 19 | 886 | 4.128 \$ |
| Mon | 6 | 80 | 388 \$ |
| | 7 | 661 | 3.099 \$ |
| | 8 | 1992 | 9.517 \$ |
| | 9 | 2482 | 11.605 \$ |
| | 10 | 2494 | 11.898 \$ |
| | 11 | 2612 | 12.427 \$ |
| | 12 | 1478 | 7.179 \$ |
| | 13 | 1308 | 6.140 \$ |
| | 14 | 1267 | 5.771 \$ |
| | 15 | 1376 | 6.463 \$ |
| | 16 | 1319 | 6.159 \$ |
| | 17 | 1355 | 6.140 \$ |
| | 18 | 1302 | 5.927 \$ |
| | 19 | 1077 | 5.006 \$ |
| Sat | 6 | 846 | 4.015 \$ |
| | 7 | 74 | 330 \$ |
| | 8 | 592 | 2.914 \$ |

Slika 6 - Tabela_4 sa atributima Month i Weekday

Peta Pivot tabela (Tabela_5) sadrži attribute iz dve tabele dataseta: Kolonu *store_location* iz tabele *Transactions*, i kolone *customer_name*, *payment_method* i *invoice_number* iz tabele *Customers*. U tabeli je prikazana količina proizvoda koju je kupio svaki kupac, kao i prihod kafeterije po svakom kupcu.

| store_location | invoice_number | payment_method | customer_name |
|-----------------|----------------|----------------|-------------------|
| Astoria | 83971 | cash | Annie Harper |
| Hell's Kitchen | 123456 | credit card | Bruce Vance |
| Lower Manhattan | 175690 | mobile | Charles Harper |
| | 182597 | | Christine Carlson |
| | 194845 | | Claude Sterling |
| | 206898 | | Courtney Evans |
| | 244981 | | Jack Fisher |
| | 260023 | | James Gates |
| | 353270 | | John Pearson |
| | 361645 | | Joseph Tilly |
| | 376021 | | Judy Waters |
| | 448329 | | Kayla Wood |
| | 479866 | | Kenneth Sanders |
| | | | Lance Willis |
| | | | Maria Johnson |
| | | | Mary Adkinson |
| | | | Michael Jenkins |
| | | | Michele Smith |

| store_location | invoice_number | payment_method | customer_name | Količina kupljenih proizvoda | Cena |
|----------------|----------------|----------------|-------------------|------------------------------|------------|
| Astoria | 694231 | credit card | Richard Bradley | 2 | 4 \$ |
| | - | - | - | 70989 | 232.240 \$ |
| | 83971 | cash | Claude Sterling | 2 | 5 \$ |
| | 194845 | mobile | Thomas Bailey | 1 | 3 \$ |
| | 206898 | credit card | James Gates | 1 | 4 \$ |
| | 361645 | credit card | Annie Harper | 1 | 3 \$ |
| | 479866 | credit card | Michele Smith | 2 | 7 \$ |
| | 557452 | credit card | Kayla Wood | 1 | 3 \$ |
| | 801501 | mobile | Torrie Tucker | 2 | 6 \$ |
| | 840758 | credit card | Christine Carlson | 2 | 6 \$ |
| | 900409 | cash | Judy Waters | 2 | 9 \$ |
| | - | - | - | 71725 | 236.473 \$ |
| | 123456 | mobile | Steve Owens | 1 | 3 \$ |
| | 175690 | cash | Theresa McMillan | 1 | 3 \$ |
| | 182597 | cash | John Pearson | 2 | 7 \$ |
| | 244981 | cash | Sanjay Gupta | 2 | 5 \$ |
| | 260023 | mobile | Richard Townsend | 1 | 4 \$ |
| | 353270 | cash | Joseph Tilly | 1 | 3 \$ |
| | 376021 | credit card | Charles Harper | 2 | 5 \$ |
| | 448329 | credit card | Bruce Vance | 2 | 7 \$ |
| | 495672 | credit card | Patrick Ewing | 1 | 4 \$ |
| | 510028 | credit card | Robert Moreland | 1 | 3 \$ |
| | 521309 | credit card | Maria Johnson | 1 | 3 \$ |
| | 581377 | mobile | Jack Fisher | 2 | 5 \$ |
| | 622584 | cash | Lance Willis | 2 | 4 \$ |
| | 752526 | mobile | Mary Adkinson | 2 | 6 \$ |
| | 768070 | mobile | Kenneth Sanders | 2 | 6 \$ |
| | 820436 | credit card | Steve James | 1 | 3 \$ |
| | 938958 | mobile | Courtney Evans | 1 | 3 \$ |
| | 965034 | cash | Michael Jenkins | 2 | 6 \$ |

Slika 7 - Tabela_5 prikazana sa kolonama

Šesta Pivot tabela sadrži dimenzije *store_location* i *Month*. U tabeli je prikazan mesečni prihod kafeterije, broj transakcija i količina prodatih proizvoda sumirani po mesecima i tri različite fizičke prodavnice kafeterija.

Month
Apr
Feb
Jan
Jun
Mar
May

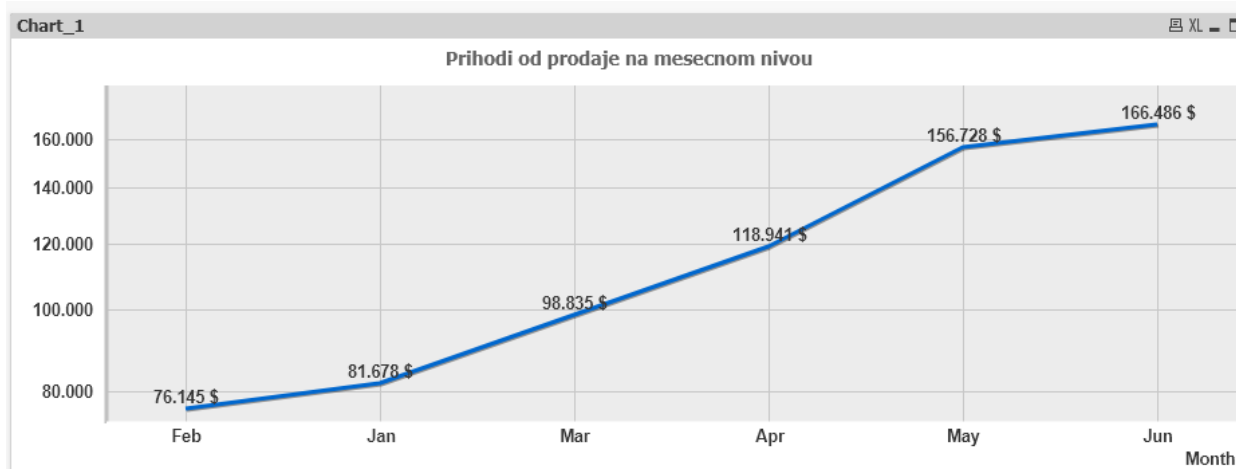
Tabela_6

| store_location | Month | Broj transakcija | Količina prodatih... | Mesečni prihod |
|-----------------|-------|------------------|----------------------|------------------|
| Astoria | Apr | 8556 | 12026 | 39477,61 |
| | Feb | 5490 | 7718 | 25105,34 |
| | Jan | 5913 | 8281 | 27313,66 |
| | Jun | 11985 | 16853 | 55083,11 |
| | Mar | 7180 | 9999 | 32835,43 |
| Hell's Kitchen | May | 11475 | 16114 | 52428,7599999999 |
| | Apr | 8629 | 12194 | 40304,14 |
| | Feb | 5606 | 7937 | 25719,8 |
| | Jan | 5868 | 8340 | 27820,65 |
| | Jun | 12071 | 17124 | 56957,08 |
| Lower Manhattan | Mar | 7267 | 10198 | 33110,57 |
| | May | 11294 | 15944 | 52598,9299999999 |
| | Apr | 8150 | 12249 | 39159,33 |
| | Feb | 5263 | 7895 | 25320,05 |
| | Jan | 5533 | 8249 | 26543,43 |
| | Jun | 11296 | 16965 | 54445,69 |
| | Mar | 6782 | 10209 | 32888,68 |
| | May | 10758 | 16175 | 51700,07 |

Slika 8 – Tabela_6

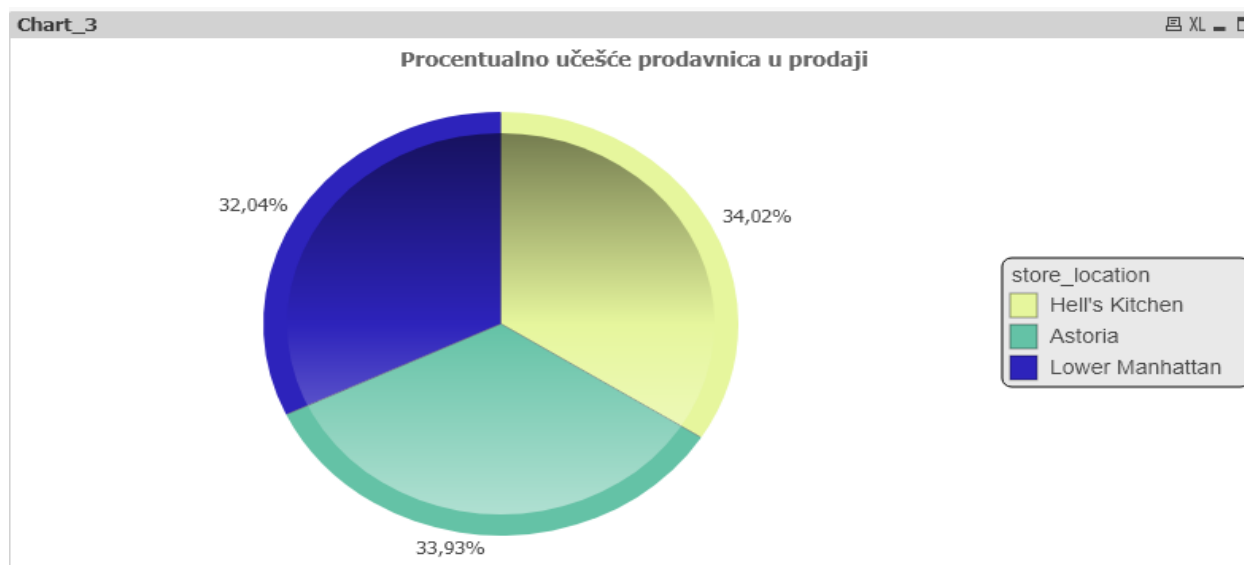
5 Kreiranje grafičkih izveštaja u QlikView softveru

Projekat sadrži ukupno 9 grafičkih prezentacija, a u tekstu će biti prikazano i opisano 8 grafika. Grafik Chart_1 prikazuje prihod od prodaje kafeterije na mesečnom nivou. Iz grafičke prezentacije se vidi da je najveći prihod zabeležen u prolećnim u letnjim mesecima (Maj i Jun), a najmanji u zimskim mesecima (Januar i Februar).



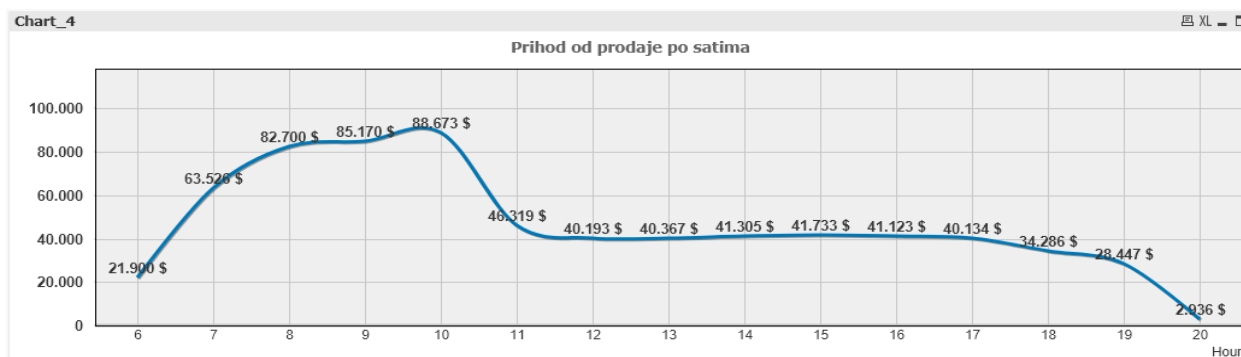
Slika 9 - Chart_1 – Prihod od prodaje na mesečnom nivou

Sledeći grafički prikaz je grafik iz kategorije Pie Chart, nazvan Chart_3, koji predstavlja udeo tri fizičke prodavnice u ukupnoj prodaji tokom posmatranog perioda od šest meseci. Na Chart_3 grafiku primećuje se da sve tri prodavnice, locirane u delovima grada Hell's Kitchen, Astoria i Lower Manhattan, imaju gotovo podjednak udeo u ukupnoj prodaji kompanije.



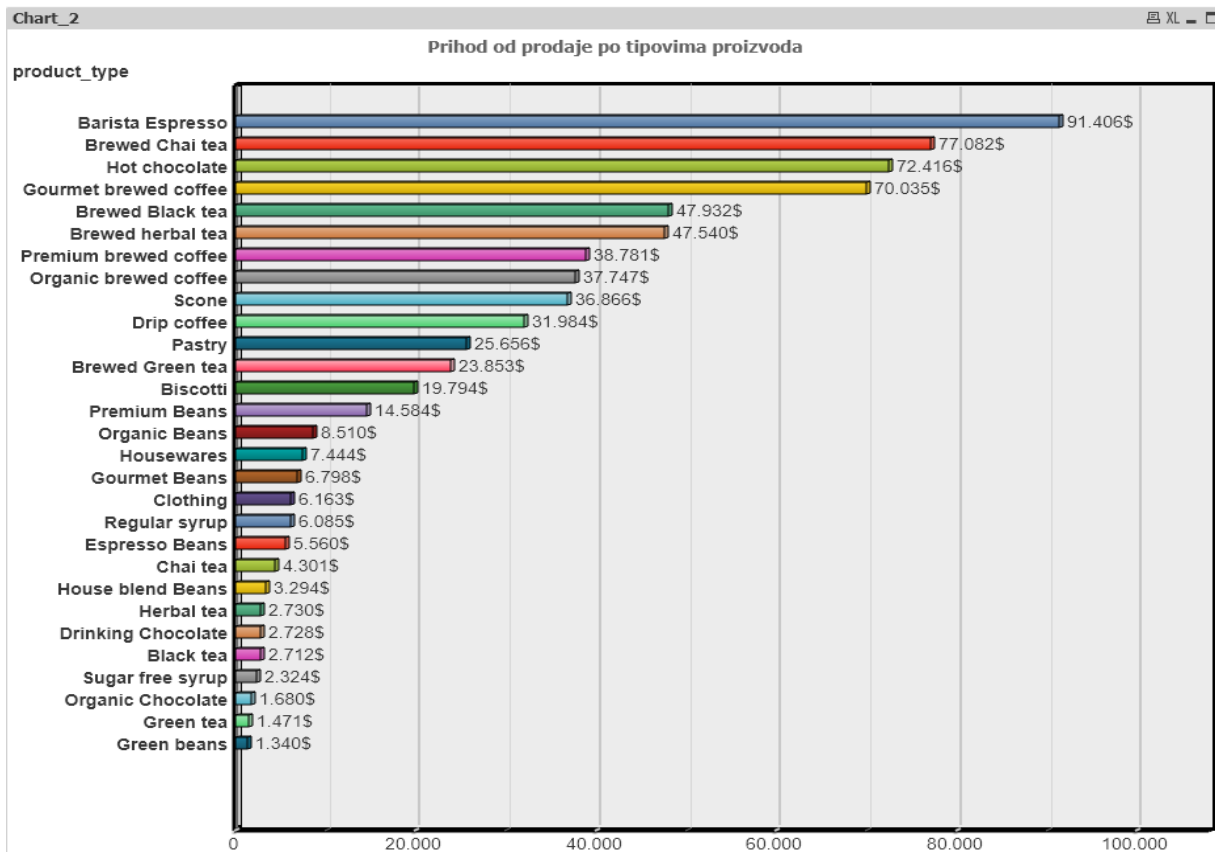
Slika 10 - Udeo u ukupnoj prodaji fizičkih prodavnica nazvanih po delovima grada u kojem se nalaze

Treći grafik je linijski dijagram Chart_4i, i predstavlja prihod od prodaje, sumiran po satima u jednom danu. Posmatrajući grafik, zaključujemo da najveći prihod ostvarujemo u jutarnjim satima, gde između 9 i 10 časova prodajemo najviše proizvoda. Kako se približavamo kraju dana, tako i prodaja opada, da bi najmanji prihod bio između 19 i 20 časova.



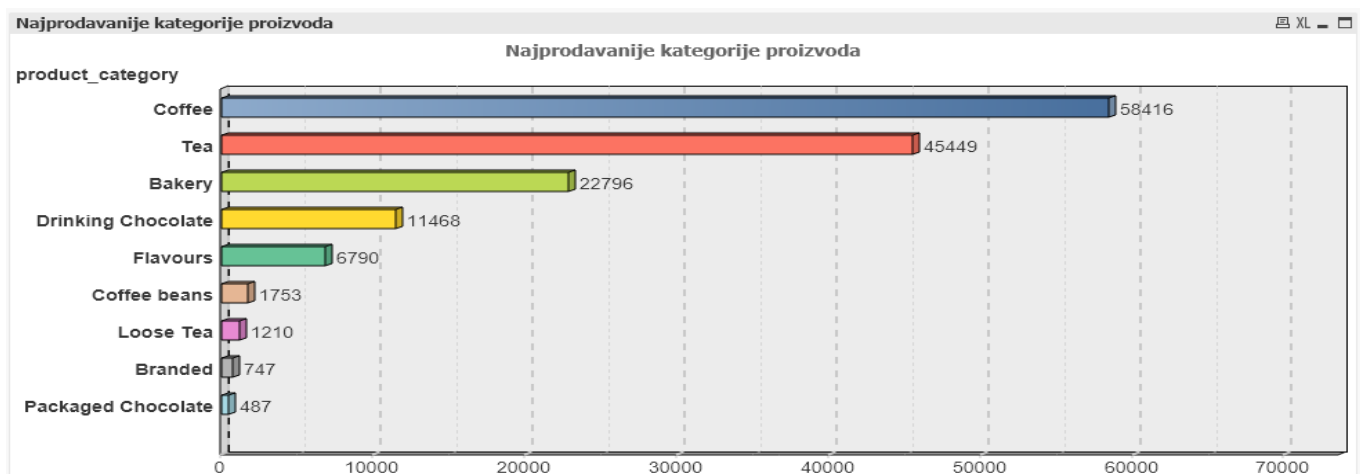
Slika 11 – Chart_4 grafik koji prikazuje prihod od prodaje u dolarima raspoređen po satima

Četvrta grafička prezentacija prikazuje najprodavanije pojedinačne proizvode, opadajuće sortirane po prihodu od prodaje koji je ostvario svaki pojedinačni proizvod, od najprodavanijih do najmanje prodvanih. Posmatrajući grafik može se zaključiti da najviše prihoda donosi prodaja različitih vrsta kafa i čajeva.



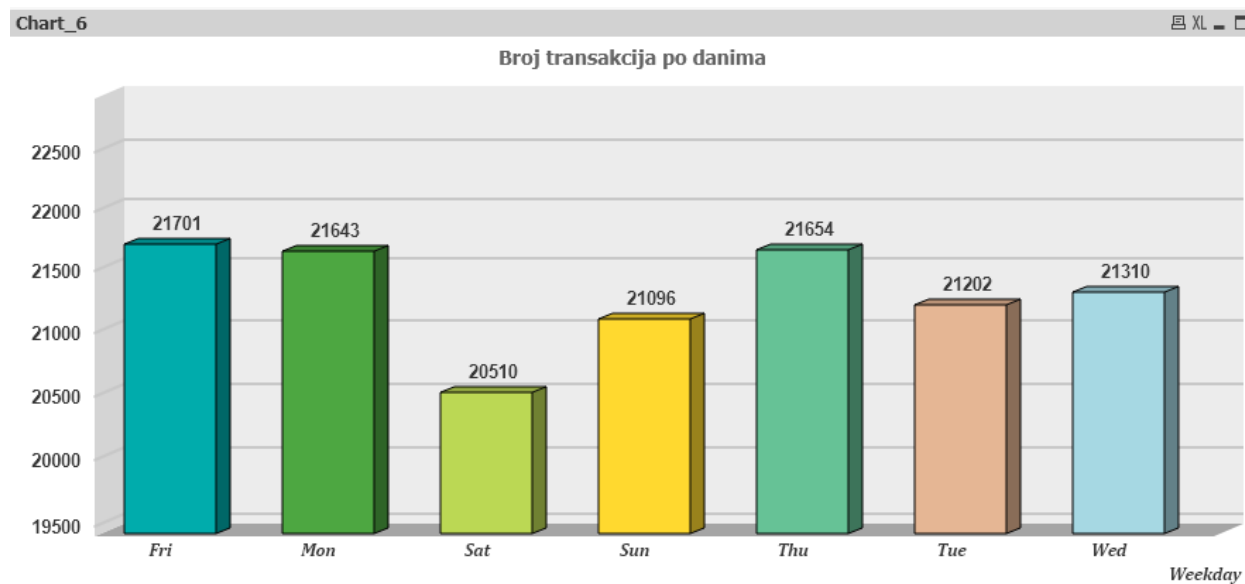
Slika 12 – Prihod od prodaje sumiran po tipu proizvoda, sortiran po vrednostima u opadaju

Peti grafik je dijagram iz kategorije Bar Chart, sa horizontalno usmerenim figurama koje prikazuju vrednosti. Na grafiku su predstavljane najprodavanije kategorije proizvoda, sa opadajućim vrednostima ukupno ostvarenog broja transakcija po kategoriji proizvoda. Najveći broj transakcija uključivao je proizvode iz kategorija Kafa i Čaj, i one predstavljaju više od polovine svih kupovina.



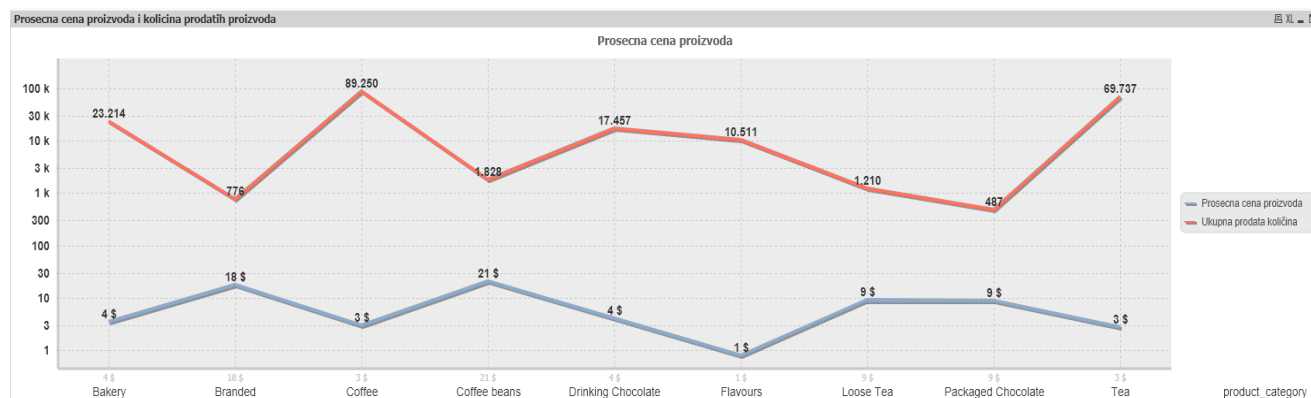
Slika 13 – Prodane količine proizvoda sumirane po kategorijama

Šesti grafik je takođe dijagram iz kategorije Bar Chart, sa vertikalno usmerenim figurama koje predstavljaju ukupan broj transakcija podeljenih po danima u nedelji. Analizirajući grafik može se zaključiti da se najviše transakcija izvrši radnim danima, a nešto manji broj kupovina odvijao se subotom i nedeljom.



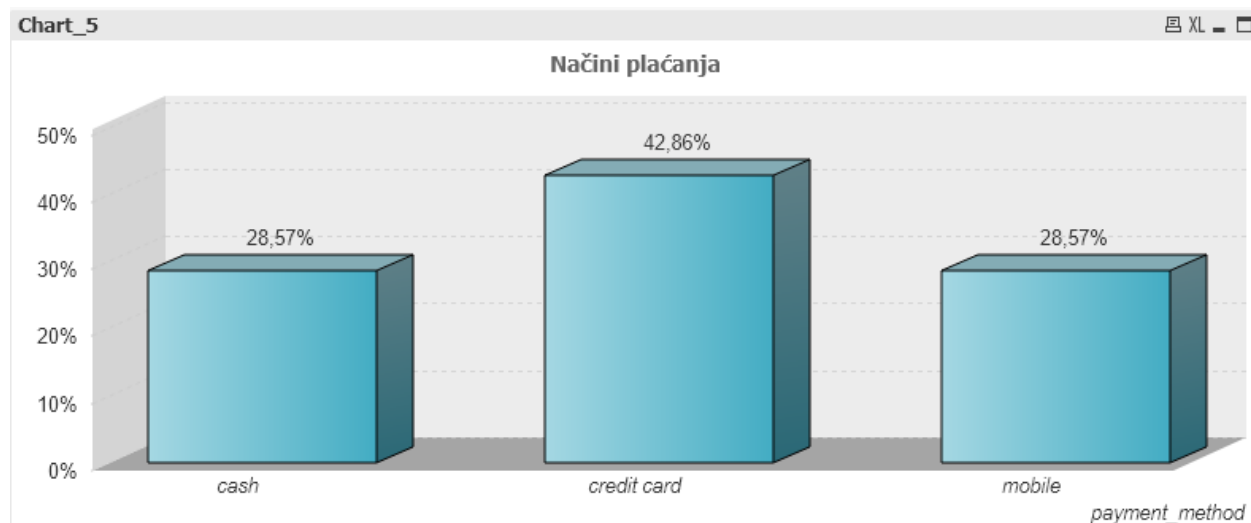
Slika 14 – Broj obavljenih kupovina po danima u nedelji

Sedma grafička prezentacija je linijski dijagram, koji prikazuje odnos između ukupne količine prodanih proizvoda po kategorijama i prosečne cene proizvoda iz odgovarajuće kategorije. Posmatrajući grafik zaključujemo da smo najviše prodali proizvoda i najveći prihod ostvarili prodajom proizvoda iz kategorije Coffee(Kafa) i Čaj(Tea), iako je prosečna cena tih proizvoda nije naročito visoka. Odlične rezultate smo ostvarili i prodajom proizvoda iz kategorije čokoladnih napitaka i pekarskih proizvoda.



Slika 15 – Prosečna cena kategorije proizvoda i količina prodanih proizvoda sumirana po kategoriji

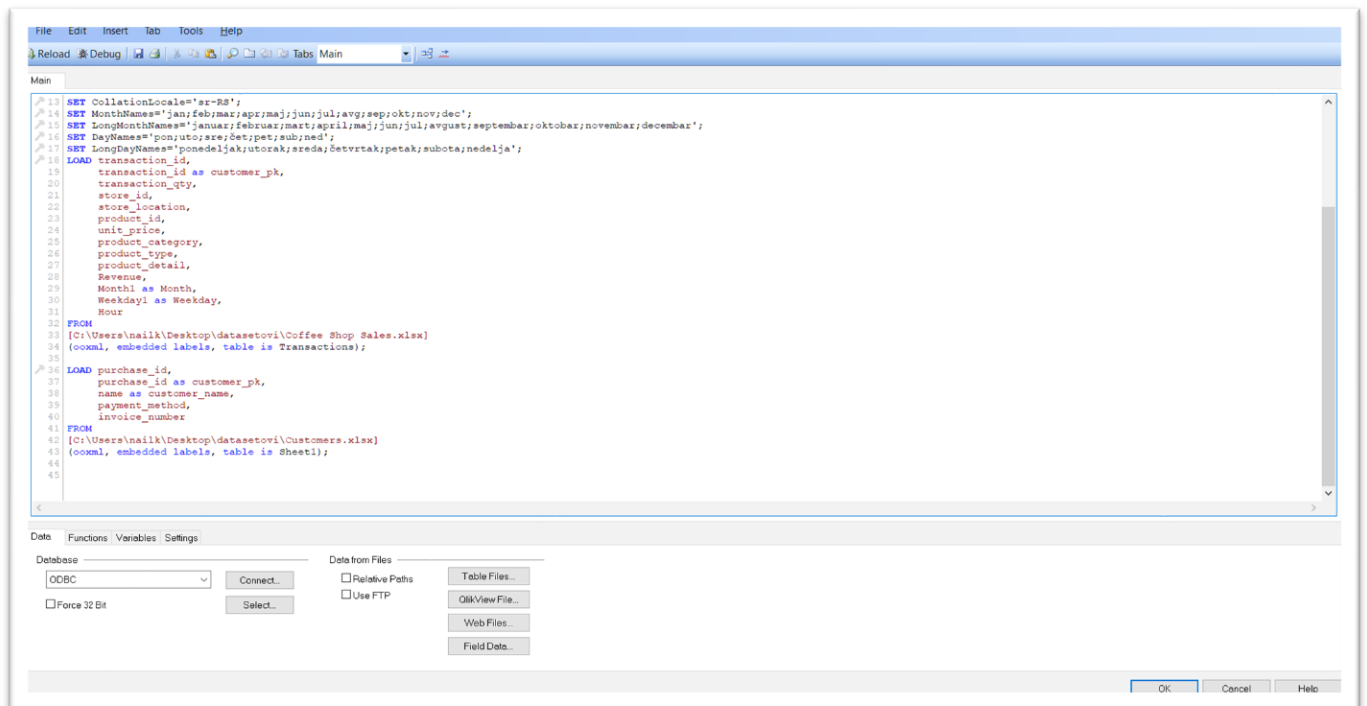
Osmi grafik je dijagram iz kategorije Bar Chart, koji predstavlja odnos između najpopularnijih načina plaćanja. Kupci su najviše koristili kreditne kartice (42,86%), a podjednako su plaćanja vršili mobilnim telefonom i gotovinom (28,57%).



Slika 16 – Najzastupljeniji načini plaćanja

6 Spajanje različitih tabela kreiranjem zajedničkog ključa u QlikView softveru

Kako bih spojio tabelu Transactions, koja sadrži detaljne podatke o kupovini proizvoda, i tabelu Customers koja sadrži podatke o kupcima, kreirao sam spoljni ključ *customer_pk*, koji spaja kolone *transaction_id* u tabeli Transactions, i kolonu *purchase_id* u tabeli Customers.



Slika 17 - Prikaz spajanje tabela Customers i Transactions pomoću spoljnog ključa customer_pk

Kako bih demonstrirao uspešno povezivanje tabela, kreirao sam Straight tabelu Tabela_pk, čije su dimenzije kolone *product_category*, *product_type* i *product_detail* iz tabele Transactions, i kolona *customer_name* iz tabele Customers. Tabela prikazuje podatke o kupovini svakog kupca, iznos njegovog računa i količinu proizvoda koju je kupio.

| product_cat... | product_type | product_detail | customer_name |
|--------------------|-----------------------|-------------------------|-------------------|
| Bakery | Barista Espresso | Almond Croissant | Annie Harper |
| Branded | Biscotti | Brazilian - Organic | Bruce Vance |
| Coffee | Black tea | Brazilian Lg | Charles Harper |
| Coffee beans | Brewed Black tea | Brazilian Rg | Christine Carlson |
| Drinking Chocolate | Brewed Chai tea | Brazilian Sm | Claude Sterling |
| Flavours | Brewed Green tea | Cappuccino | Courtney Evans |
| Loose Tea | Brewed herbal tea | Cappuccino Lg | Jack Fisher |
| Packaged Chocolate | Chai tea | Carmel syrup | James Gates |
| Tea | Clothing | Chili Mayan | John Pearson |
| | Drinking Chocolate | Chocolate Chip Biscotti | Joseph Tilly |
| | Drip coffee | Chocolate Croissant | Judy Waters |
| | Espresso Beans | Chocolate syrup | Kayla Wood |
| | Gourmet Beans | | Kenneth Sanders |
| | Gourmet brewed coffee | | Lance Willis |
| | Green beans | | Maria Johnson |
| | | | Mary Adkinson |
| | | | Michael Jenkins |
| | | | Michele Smith |
| | | | Patrick Ewing |
| | | | Richard Bradley |
| | | | Richard Townsend |
| | | | Robert Moreland |
| | | | Sanjay Gupta |
| | | | Steve James |
| | | | Steve Owens |

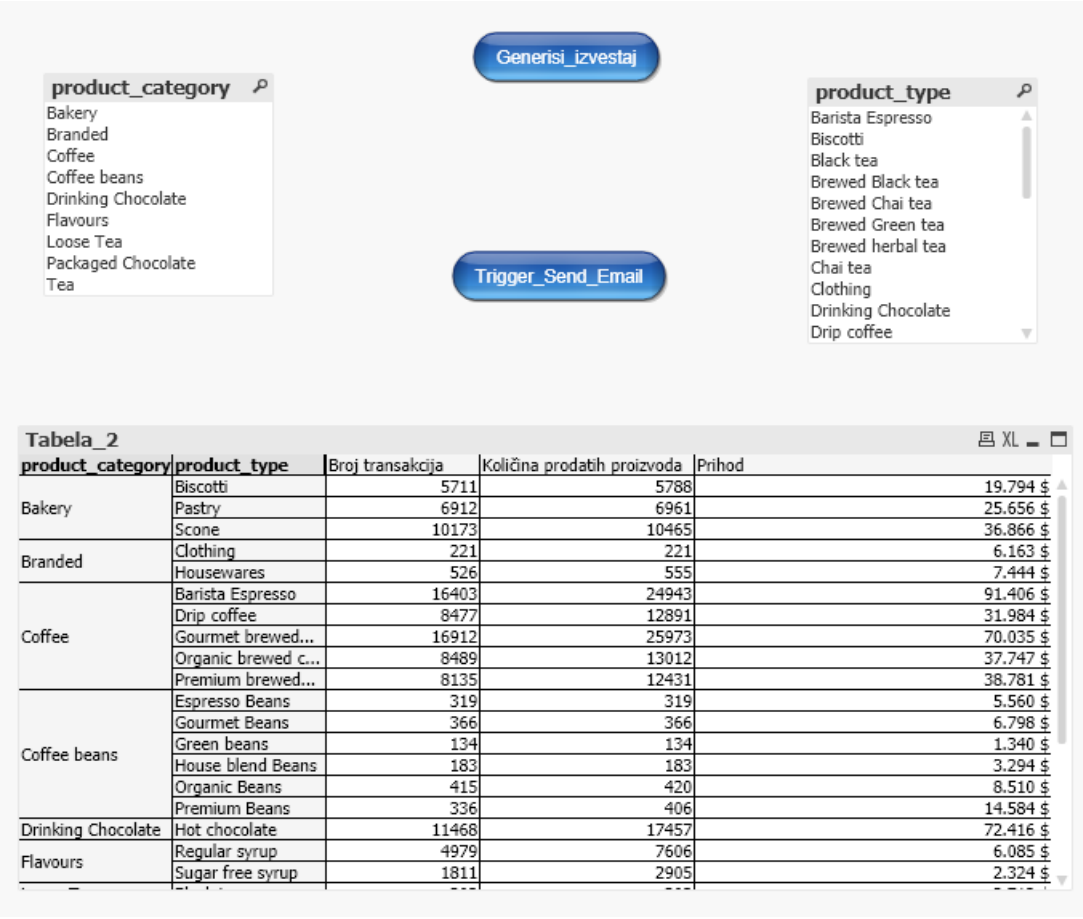
| Tabela_pk | | | | | |
|------------------|---------------------|---------------------|-------------------|-----------|----------|
| product_category | product_type | product_detail | customer_na... | Račun(\$) | Količina |
| Tea | Brewed herbal tea | Lemon Grass Rg | Annie Harper | 2,5 | 1 |
| Coffee | Organic brewed c... | Brazilian Lg | Bruce Vance | 7 | 2 |
| Tea | Brewed Chai tea | Spicy Eye Opener... | Charles Harper | 5,1 | 2 |
| Tea | Brewed Black tea | Earl Grey Lg | Christine Carlson | 6 | 2 |
| Tea | Brewed herbal tea | Peppermint Rg | Claude Sterling | 5 | 2 |
| Tea | Brewed Chai tea | Spicy Eye Opener... | Courtney Evans | 2,55 | 1 |
| Tea | Brewed Chai tea | Morning Sunrise... | Jack Fisher | 5 | 2 |
| Coffee | Barista Espresso | Cappuccino | James Gates | 3,75 | 1 |
| Coffee | Gourmet brewed... | Ethiopia Lg | John Pearson | 7 | 2 |
| Coffee | Drip coffee | Our Old Time Din... | Joseph Tilly | 3 | 1 |
| Coffee | Barista Espresso | Cappuccino Lg | Judy Waters | 8,5 | 2 |
| Tea | Brewed Chai tea | Morning Sunrise... | Kayla Wood | 2,5 | 1 |
| Tea | Brewed Green tea | Serenity Green T... | Kenneth Sanders | 6 | 2 |
| Coffee | Organic brewed c... | Brazilian Sm | Lance Willis | 4,4 | 2 |
| Bakery | Scone | Ginger Scone | Maria Johnson | 3,25 | 1 |
| Tea | Brewed Green tea | Serenity Green T... | Mary Adkinson | 6 | 2 |
| Tea | Brewed herbal tea | Lemon Grass Lg | Michael Jenkins | 6 | 2 |
| Coffee | Gourmet brewed... | Ethiopia Lg | Michele Smith | 7 | 2 |
| Bakery | Biscotti | Ginger Biscotti | Patrick Ewing | 3,5 | 1 |
| Coffee | Organic brewed c... | Brazilian Sm | Richard Bradley | 4,4 | 2 |
| Bakery | Biscotti | Ginger Biscotti | Richard Townsend | 3,5 | 1 |
| Coffee | Organic brewed c... | Brazilian Rg | Robert Moreland | 3 | 1 |
| Tea | Brewed Chai tea | Spicy Eye Opener... | Sanjay Gupta | 5,1 | 2 |
| Coffee | Drip coffee | Our Old Time Din... | Steve James | 2,5 | 1 |
| Tea | Brewed Chai tea | Spicy Eye Opener... | Steve Owens | 2,55 | 1 |
| | | | | ##### | 214470 |

Slika 18 – Tabela_pk koja prikazuje vrednosti iz dve tabele Customers i Transactions

Uvidom u tabele Transactions i Customers, pokazuje se da je povezivanje tabela preko kolona *transaction_id* i *purchase_id* spoljnim ključem *customer_pk* uspešno izvršeno, iz razloga što Tabela_pk u jednom redu prikazuje vrednosti iz tabele Customers, i iz tabele Transactions, tamo gde kolone *transaction_id* i *purchase_id* imaju identičnu vrednost.

7 Automatsko slanje izveštaja sa novogenerisanim podacima mejlom, svakog dana

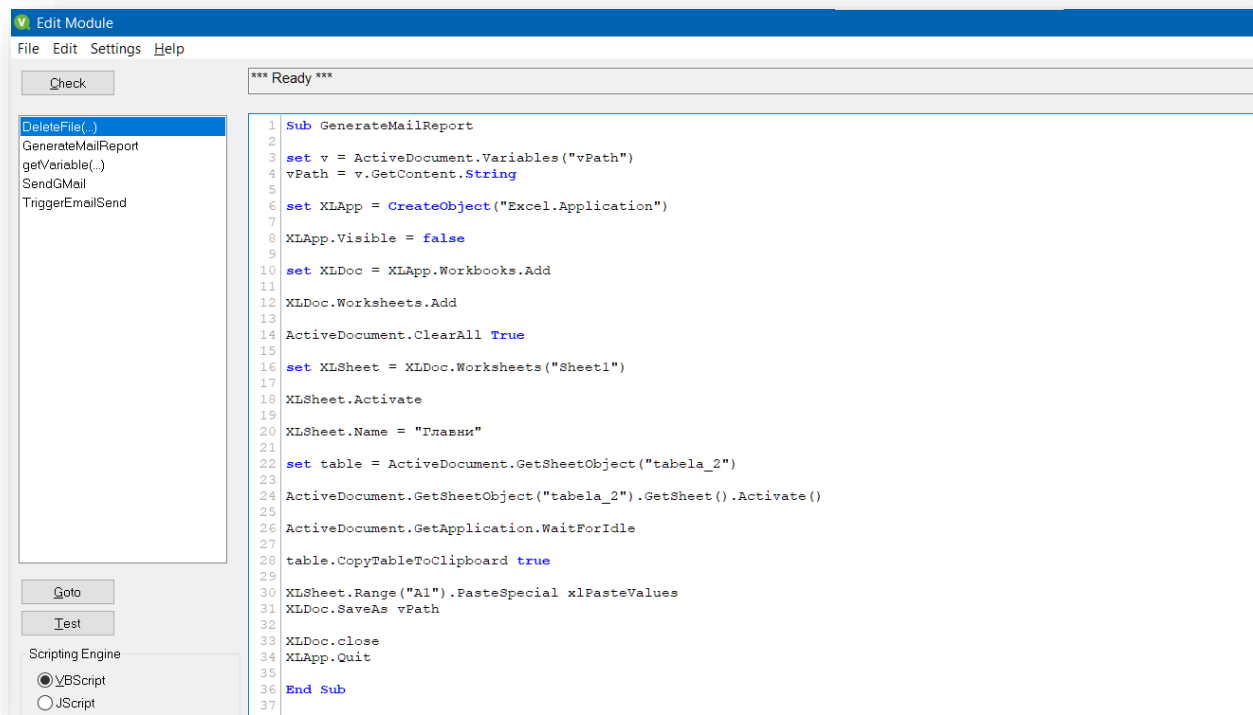
Procedura automatskog slanja izveštaja e-mailom sastoji se iz nekoliko koraka. Prvo, biramo tabelu, grafik ili neki drugi objekat iz softvera QlikView, na osnovu kojeg će se generisati izveštaj. Za potrebe projekta, kreiraće se izveštaj od tabele *Tabela_2*. U *Edit Module* prozoru u QlikViewu pišemo manji program, odnosno subroutine-u ili Sub u Visual Basic programskom jeziku. Prvi mali program nosi ime *GenerateMailReport*. Da bi program pravilno funkcionisao, moramo definisati promenljivu *vPath* i dodeliti joj vrednost adrese foldera našeg sistema u kojem će se izveštaj generisati. *vPath* promenljivu definišemo tako što u QlikView izaberemo putanju *Document Properties -> Variables*. Klikom na dugme New, kreiramo novu promenljivu i definišemo ime i lokaciju u našoj sistemskoj hijerarhiji gde će se izveštaj pojaviti. Nakon pisanja i provere programskog koda u prozoru *Edit Module*, kreiramo objekat Button pod imenom *Generisi_izvestaj* u QlikViewu kako bi pokrenuli postupak kreiranja izveštaja. Ukoliko je procedura dobro napisana, izveštaj u .xlsx formatu pojaviće se u našem sistemu na lokaciji koju smo naznačili u QlikView programu. Proces kreiranja izveštaja pokrećemo klikom na dugme *Generisi_izvestaj*. Ukoliko je subroutine-a napisana pravilno i promenljiva *vPath* tačno definisana, Excel fajl sa imenom Izveštaj pojaviće se na našem sistemu.



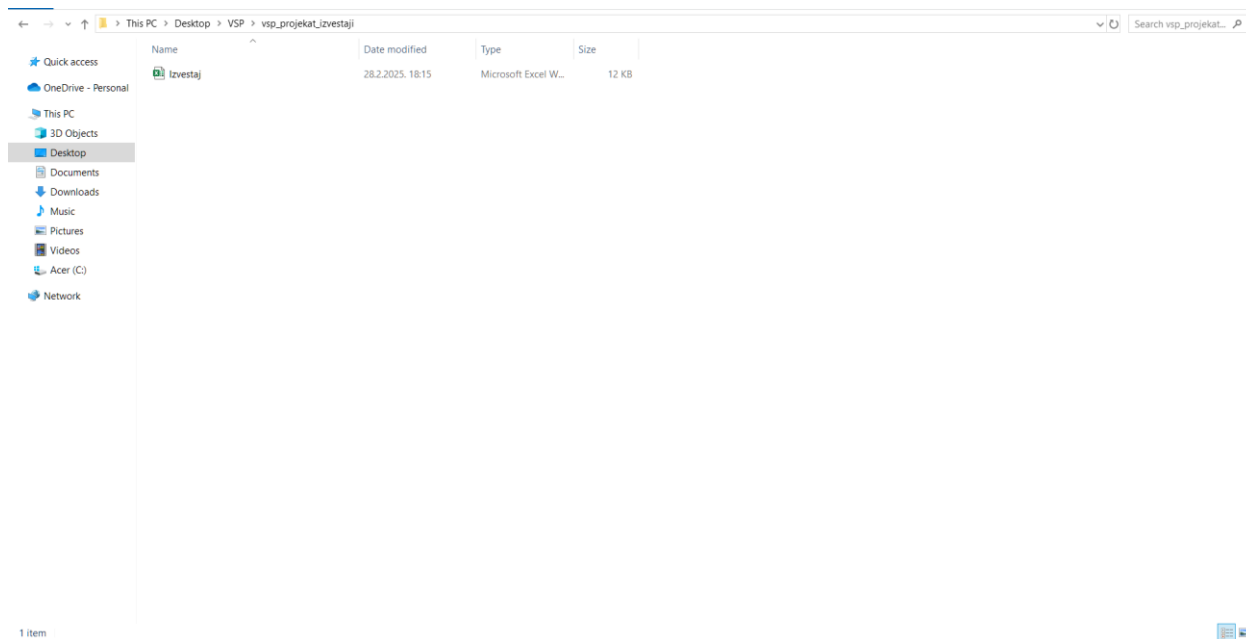
The screenshot shows the QlikView interface. At the top, there are two dropdown menus: 'product_category' and 'product_type'. The 'product_category' menu is open, showing options like Bakery, Branded, Coffee, etc. The 'product_type' menu is also open, showing options like Barista Espresso, Biscotti, Black tea, etc. Below these menus are two buttons: 'Generisi_izvestaj' and 'Trigger_Send_Email'. At the bottom, there is a table titled 'Tabela_2' with the following data:

| product_category | product_type | Broj transakcija | Količina prodanih proizvoda | Prihod |
|--------------------|---------------------|------------------|-----------------------------|-----------|
| Bakery | Biscotti | 5711 | 5788 | 19.794 \$ |
| | Pastry | 6912 | 6961 | 25.656 \$ |
| | Scone | 10173 | 10465 | 36.866 \$ |
| Branded | Clothing | 221 | 221 | 6.163 \$ |
| | Housewares | 526 | 555 | 7.444 \$ |
| Coffee | Barista Espresso | 16403 | 24943 | 91.406 \$ |
| | Drip coffee | 8477 | 12891 | 31.984 \$ |
| | Gourmet brewed... | 16912 | 25973 | 70.035 \$ |
| | Organic brewed c... | 8489 | 13012 | 37.747 \$ |
| | Premium brewed... | 8135 | 12431 | 38.781 \$ |
| Coffee beans | Espresso Beans | 319 | 319 | 5.560 \$ |
| | Gourmet Beans | 366 | 366 | 6.798 \$ |
| | Green beans | 134 | 134 | 1.340 \$ |
| | House blend Beans | 183 | 183 | 3.294 \$ |
| | Organic Beans | 415 | 420 | 8.510 \$ |
| | Premium Beans | 336 | 406 | 14.584 \$ |
| Drinking Chocolate | Hot chocolate | 11468 | 17457 | 72.416 \$ |
| Flavours | Regular syrup | 4979 | 7606 | 6.085 \$ |
| | Sugar free syrup | 1811 | 2905 | 2.324 \$ |

Slika 19 – Tabela_2 na osnovu koje kreiramo izveštaj



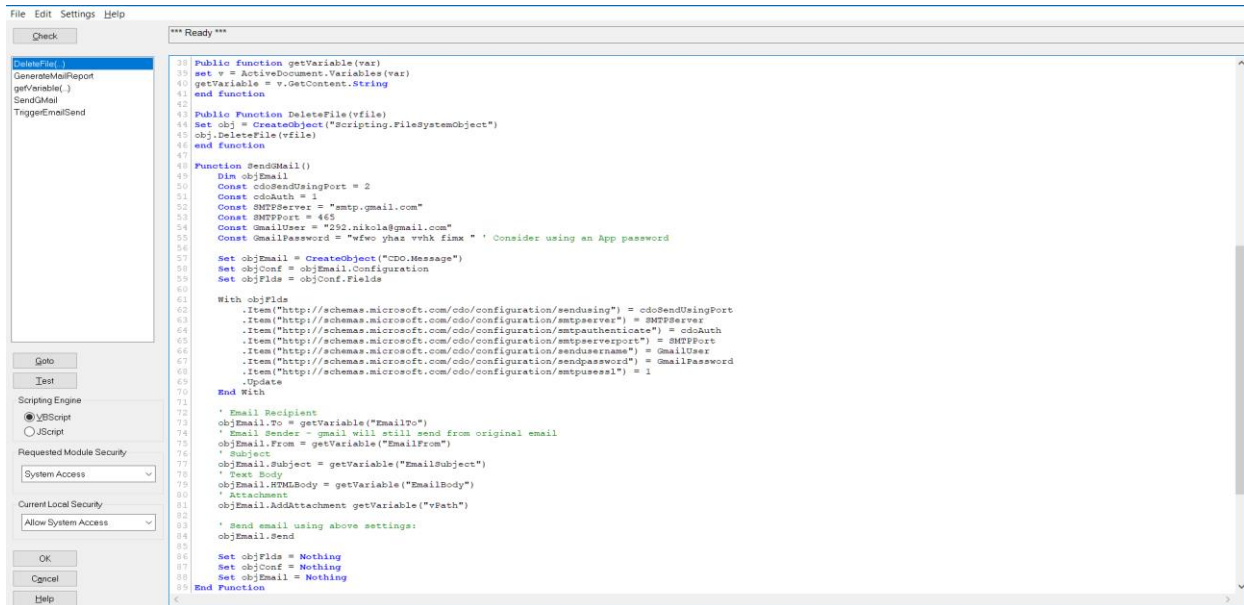
Slika 20 – GenerateMailReport subroutine-a



Slika 21 – Uspešno kreiran Excel fajl Izveštaj

Pošto je procedura kreiranja izveštaja uspešno završena, sledeći korak je pisanje novog manjeg programa za automatsko slanje izveštaja sa jedne e-mail adrese na drugu. Sam program je malo komplikovaniji od prethodnog i sastoji se iz nekoliko koraka: generisanje .xlsx izveštaja, definisanje e-mail adrese sa koje će izveštaj biti poslat, unos e-mail adrese na koju će izveštaj stići, definisanje teme i sadržaja maila i pisanje procedure za automatsko brisanje izveštaja nakon slanja.

Pre svega, na već defisanu `GenerateMailReport` subroutine-u, dodajemo dva manja potprograma: `SendGMail` i `TriggerEmailSend`.

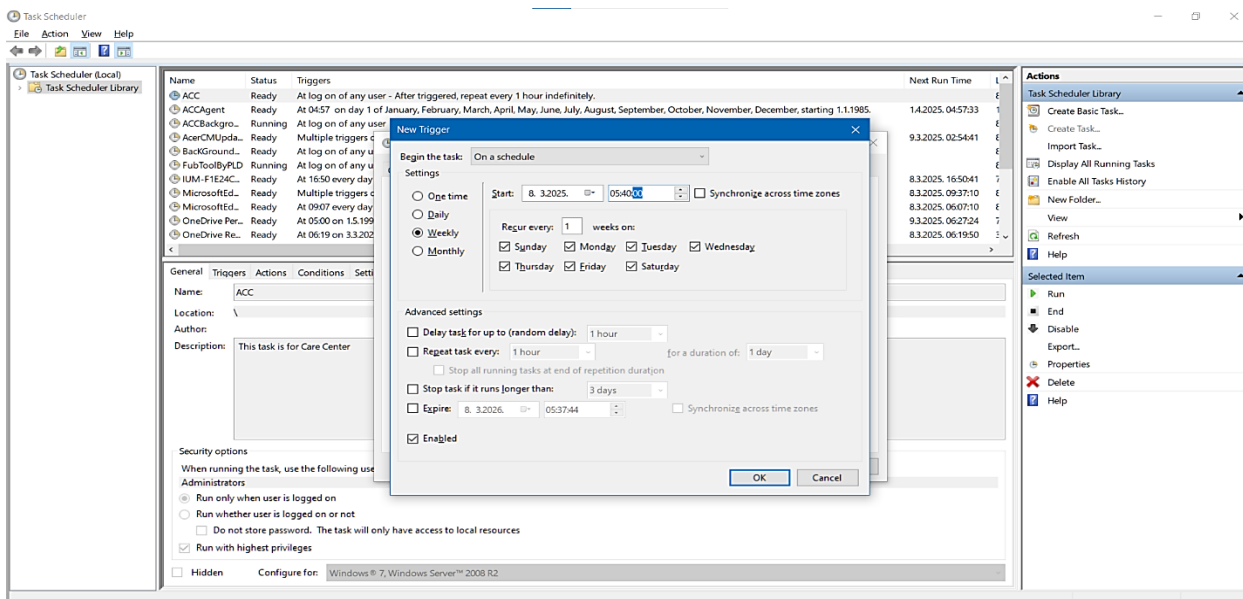


Slika 22 – Program za generisanje i automatsko slanje izveštaja na mail

Nakon što smo napisali kompletan program i testirali njegovu funkcionalnost, u sekciji *Document Properties* -> *Variables* definišemo četiri nove promenljive:

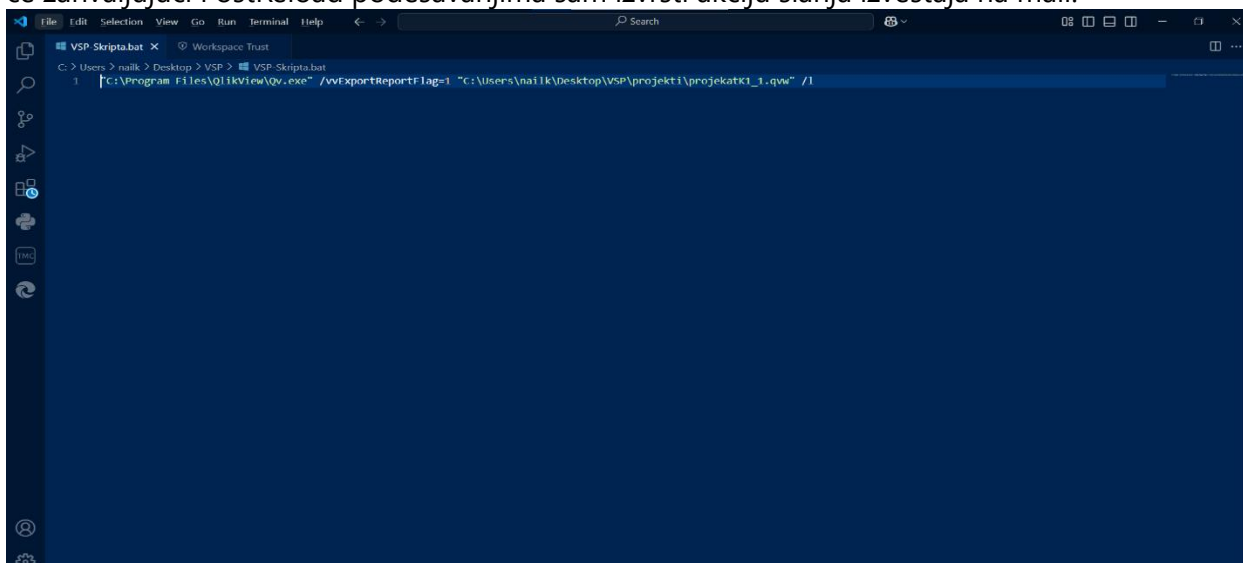
- EmailFrom – Gmail adresa sa koje šalјemo prethodno kreirani izveštaj
- EmailTo – Email adresa na koju šalјemo izveštaj
- EmailSubject – tema poslatog emaila
- EmailBody – sadržaj email poruke

Nakon što definišemo nove promenljive, QlikView program podešavamo tako da pri automatskom otvaranju QlikView projektnog filea pokrene subroutine-u za slanje izveštaja na mail. U putanji *Document Properties* -> *Triggers* pronalazimo *PostReload* akciju i dodeljuјemo joj *TriggerEmailSend* subroutine-u, i tako nalažemo QlikViewu da nakon otvaranja sam pokrene (triggeruje) proceduru automatskog slanja izveštaja na mail.



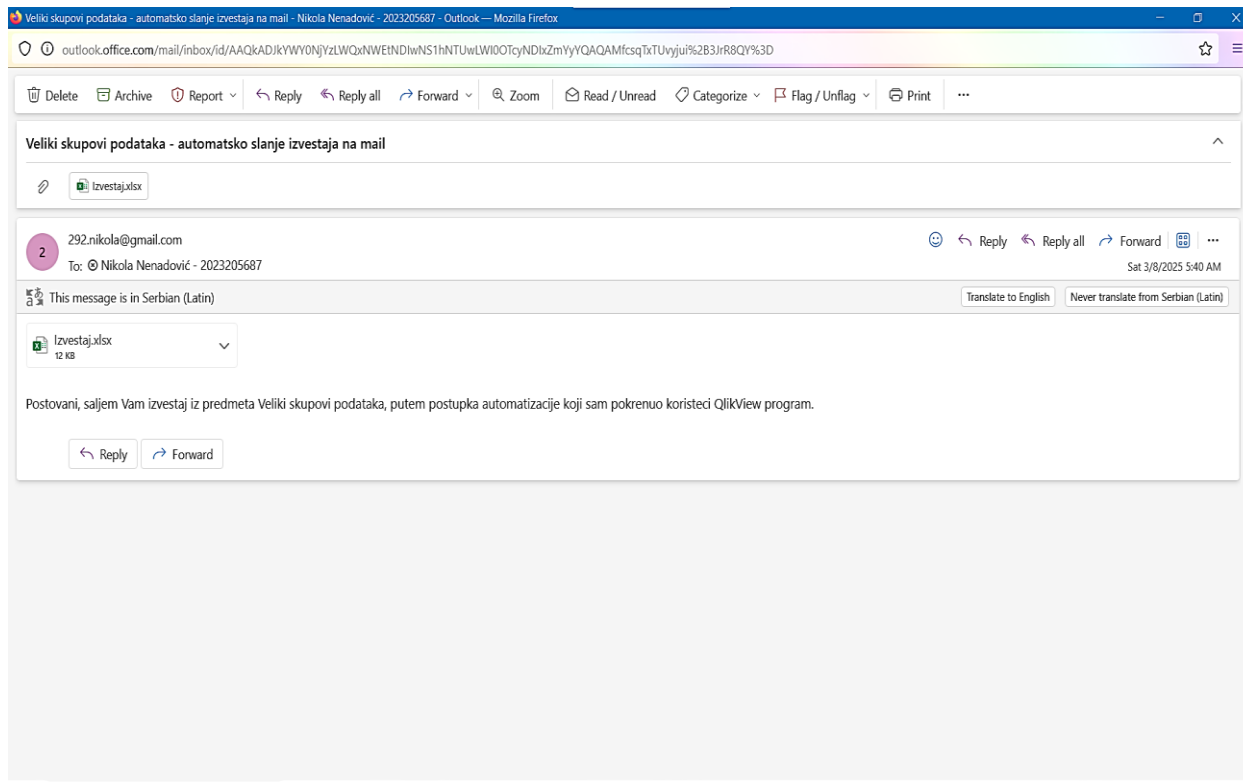
Slika 23 - Kreiranje taska za automatsko slanje izveštaja na mail svakog dana

Pošto smo završili podešavanja QlikView programa, U Windows sistemu otvaramo Task Scheduler program u kojem želimo da konfiguršemo automatsko pokretanje procedure slanja izveštaja na mail svakog dana. Nakon pokretanja Task Schedulera, u desnom uglu biramo opciju Create Task, sa ciljem kreiranja taska za slanje maila. Task ćemo imenovati Automatsko slanje izveštaja na mail. U sekciji Triggers, kreiramo novi trigger u kojem definišemo tačno vreme pokretanja taska (svakog dana, sata, nedelje ili meseca, itd.), koje je u našem slučaju svakog dana u 05:40 časova. U sekciji Actions, potrebno je da definisemo lokaciju skriptnog fajla koji će se automatski pokrenuti svakog dana u vreme koje smo naveli, i koji će nakon pokretanja naložiti Windows da otvori naš Qlikview projektni fajl. Nakon što se projektni fajl automatski otvori zahvaljujući komandama iz skriptnog fajla, Qlikview projekat će zahvaljujući PostReload podešavanjima sam izvršiti akciju slanja izveštaja na mail.



Slika 24 – Skriptni .bat fajl koji ce Task Scheduler pokrenuti svakog dana u naznačeno vreme

Adresa na koju izveštaj stiže je definisana u promenljivoj EmailTo, i to je studentska Singimail adresa. Nakon što otvorimo nalog, videćemo mail poruku koja je pristigla u 5 časova i 40 minuta, što je tačno ono vreme koje smo definisali u task za automatsko slanje mail-a. Tema i sadržaj mail poruke su identični vrednosti promenljivih EmailSubject i EmailBody, a u attachemntu maila nalazi se i sam izveštaj.



Slika 25 – Singimail interfejs koji prikazuje da je e-mail poruka zajedno sa izveštajem stigla u 05:40

Kada otvorimo Outlook servis i pristupimo Singimail adresi, vidi se da je mail sa izveštajem stigao tačno i 05 i 40 časova. Očekuje se da e-mail poruka stiže svakog dana u isto vreme, zahvaljujući automatizaciji procedure slanja koju smo sproveli koristeći Qlikview i Task Scheduler. Kada otvorimo attachment maila, prikazaće nam se izveštaj u Excel Workobook formatu.

Slika 26 – Excel izveštaj prikazan unutar Outlook interfejsa