

2023

Integrasi Dan Visualisasi Data Untuk Peningkatan Pengalaman Belanja Online

Data Crafters

Ketua Tim: Yusup Ibrahim Nursidiq

Anggota: Egi Cheria Abadi Pelawi

Deni Supriadi

Widia Febriyola



INTRODUCTION

Integrasi Dan Visualisasi Data Untuk Peningkatan Pengalaman Belanja Online

Dengan integrasi data dari berbagai platform e-commerce, tentang produk terkadang menjadi tersebar dan sulit diakses. Di sisi lain, pengguna sering mengalami ketidaknyamanan dalam pelayanan pengiriman product. Melalui integrasi data yang cerdas dan penggunaan visualisasi yang efektif, project ini bertujuan meningkatkan pengalaman Dengan penekanan pada pengalaman pengguna yang lebih baik, project ini berupaya untuk meningkatkan kualitas pelayanan.

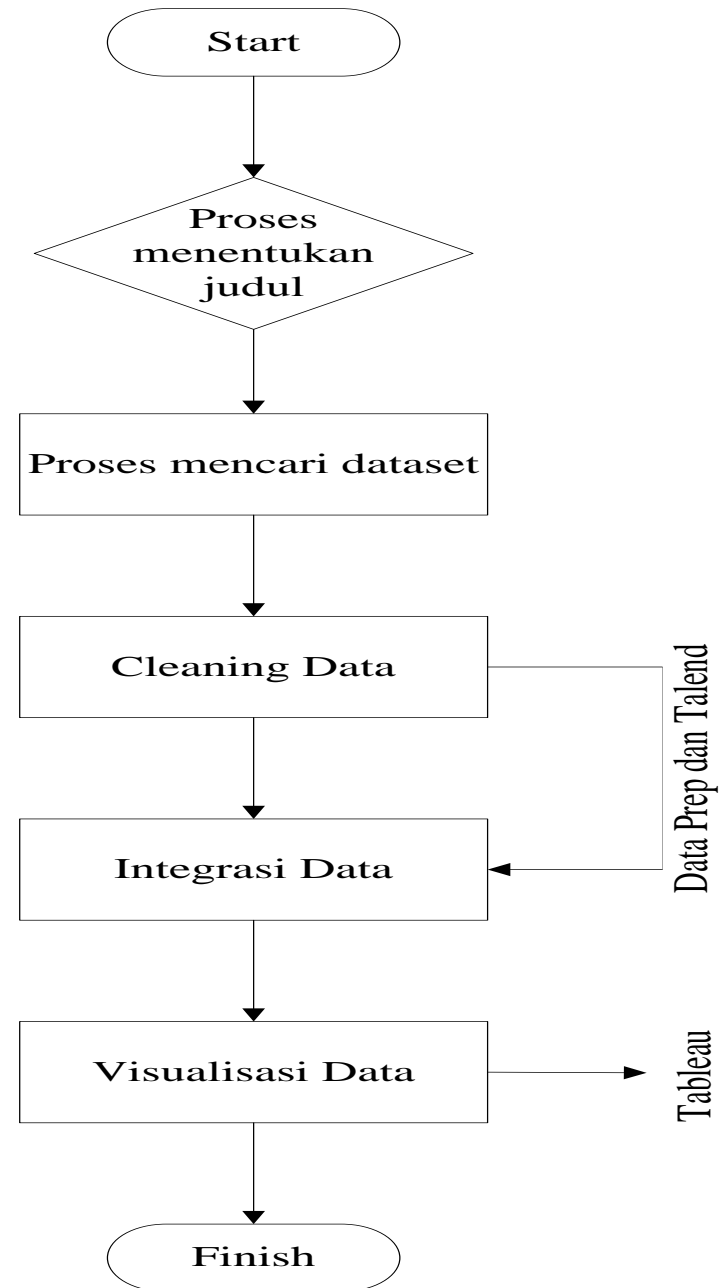




Mencari ***Data*** Melalui



Flowchart





Proses Evaluasi Data

Definisi Cleansing Data

Proses cleaning data atau pembersihan dilakukan terhadap data yang redundan atau ganda, inkonsisten, missing value dan outlier data. Hal tersebut perlu dilakukan agar tidak mempengaruhi performa proses klasifikasi yang dilakukan. Pembersihan data yang dilakukan juga menghapus fitur yang tidak diperlukan dalam proses klasifikasi nantinya.

Menurut para ahli:

1. Pembersihan data adalah operasi yang dilakukan pada data untuk menghilangkan anomali dan memperoleh data koleksi yang merupakan representasi akurat dan unik dari dunia mini. (Müller, Heiko, and Johann-Christoph Freytag. (2003))
2. Hadley Wickham, seorang pakar analisis data dan pengembang perangkat lunak statistik, menyatakan bahwa "Cleaning data adalah pekerjaan yang tidak pernah selesai." Wickham menekankan bahwa membersihkan data adalah langkah yang perlu dilakukan secara berulang karena data dapat terus berubah dan mengalami perubahan struktur.

Berikut beberapa proses yang kita lakukan pada cleaning data:

Proses Evaluasi Data

1. Proses Cleansing Data dengan Dataprep

- 1 Delete rows with missing values in reviewTitle
- 2 Convert text in name to Propercase
- 3 Set reviewContent to IFMISSING(\$col, "")
- 4 Convert text in reviewContent to Propercase
- 5 Convert text in reviewTitle to Propercase
- 6 Rename itemId to 'ItemID'
- 7 Rename category to 'Category'
- 8 Rename name to 'Name'
- 9 Rename rating to 'Rating'
- 10 Rename reviewTitle to 'ReviewTitle'
- 11 Rename reviewContent to 'ReviewContent'
- 12 Rename likeCount to 'LikeCount'
- 13 Rename upVotes to 'UpVotes'
- 14 Rename downVotes to 'DownVotes'
- 15 Rename helpful to 'Helpful'
- 16 Rename relevanceScore to 'RelevanceScore'

- 17 Rename boughtDate to 'BoughtDate'
- 18 Rename clientType to 'ClientType'
- 19 Rename retrievedDate to 'RetrievedDate'
- 20 Delete rows with missing values in BoughtDate
- 21 Delete rows with mismatched values in BoughtDate
- 22 Convert text in ClientType to Propercase
- 23 Create dateFormat_RetrievedDate from Change date format of RetrievedDate to MMM-d-yyyy
- 24 Standardize ReviewContent
- 25 Remove symbols from ReviewContent
- 26 Remove symbols from ReviewTitle
- 27 Create dateFormat_RetrievedDate1 from Change date format of RetrievedDate to MMM-d-yyyy
- 28 Rename dateFormat_RetrievedDate1 to 'RetrievedDatee'
- 29 Delete rows with missing values in ReviewTitle
- 30 Delete rows with missing values in ReviewContent

- 31 Delete originalRating
- 32 Delete RetrievedDate
- 33 Rename RetrievedDatee to 'RetrievedDate'
- 34 Delete dateFormat_RetrievedDate
- 35 Remove symbols from Category
- 36 Standardize Category
- 37 Keep rows where (1 <= LikeCount) && (LikeCount < 20)
- 38 Remove symbols from ClientType



-
- 1 **Set** Gender to IFMISSING(\$col, 'F')
 - 2 **Set** CustomerID to IFMISSING(\$col, 0)
 - 3 **Set** Coupon_Code to IFMISSING(\$col, "")
 - 4 **Set** Discount_pct to IFMISSING(\$col, 20)
 - 5 **Set** Coupon_Code to IFMISSING(\$col, 'SJ10')





Definisi Integrasi Data

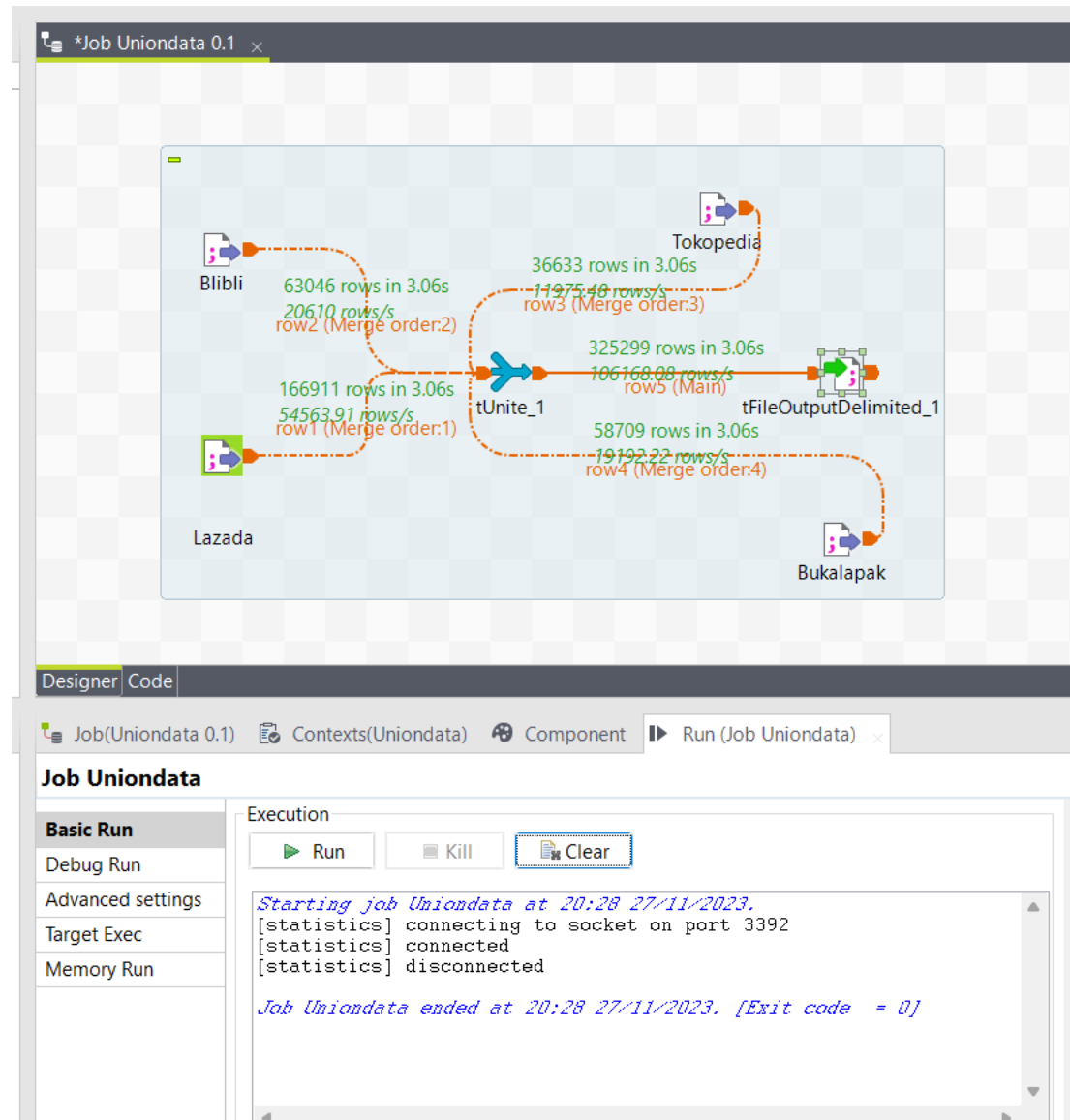
Integrasi data adalah proses menggabungkan data yang berasal dari berbagai sumber dalam suatu organisasi. Tujuannya adalah untuk membangun kumpulan data yang komprehensif, tepat, dan terkini yang berfungsi sebagai landasan bagi intelijen bisnis (BI), analisis data, dan aplikasi serta prosedur operasional penting lainnya. Replikasi, penyerapan, dan transformasi data dilakukan untuk menggabungkan berbagai jenis data ke dalam format standar yang dapat disimpan di gudang data atau data lake.

Melansir TIBCO, data integration adalah suatu proses kompilasi data dari berbagai sumber untuk memberikan user satu entitas kerja terpusat. Lalu, Safe Software menambahkan bahwa seluruh data dengan berbagai format dan tipe tersebut nantinya dikumpulkan di suatu lokasi bernama data warehouse atau gudang data.

Dalam konteks bisnis dan teknologi informasi, integrasi data sering kali mencakup penerapan teknologi ETL (Extract, Transform, Load) dan penggunaan alat atau platform integrasi data untuk memfasilitasi proses tersebut.

Berikut merupakan proses Integrasi Data yang kita lakukan:

2. Proses Integrasi Data





Definisi Visualisasi Data

Visualisasi data adalah representasi grafis dari informasi dan data. Tujuannya adalah untuk membantu orang memahami konteks, menganalisis pola, dan membuat keputusan berdasarkan informasi yang disajikan secara visual.

Definisi visualisasi data dari beberapa para ahli sebagai berikut:

Edward Tufte: Edward Tufte, seorang pakar visualisasi data terkenal, menggambarkan visualisasi data sebagai "transformasi kesejajaran ide-ide kognitif menjadi tanda-tanda grafis."

Hadley Wickham: Hadley Wickham, seorang ahli analisis data, menyatakan bahwa visualisasi data adalah "cara menyajikan data sehingga mudah dimengerti, mengundang eksplorasi, dan mengekspresikan pola dalam data."

Berikut merupakan beberapa proses visualisasi yang kami lakukan di Tableau:

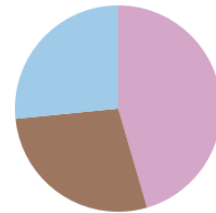


1. Pengaruh Kupon Terhadap Pembelian Customer

TOP 10 Popular Coupon Code September

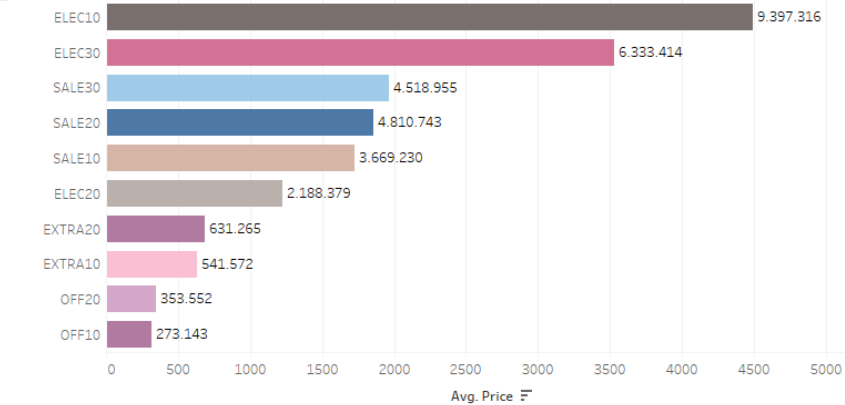
Month of Tr...

September



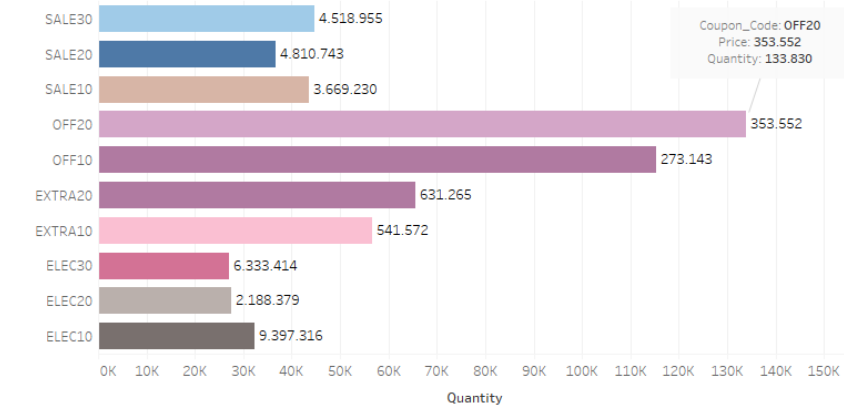
Coupon Code By Quantity With AVG Price Top 10

Coupon_Co..

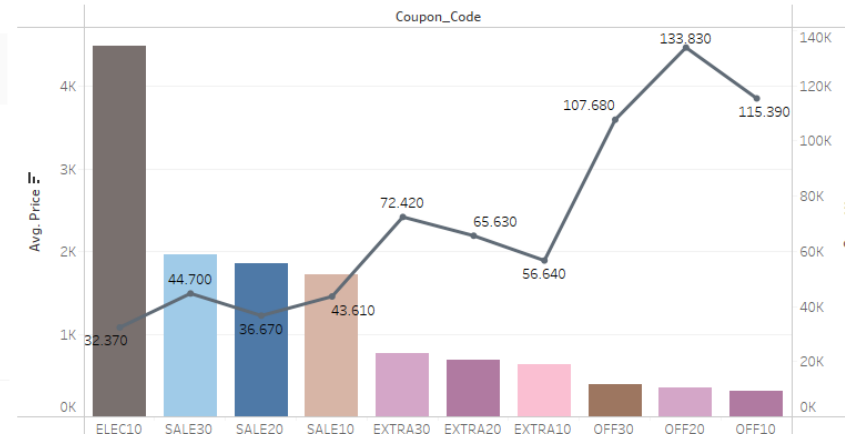


Coupon Code By Quantity With AVG Quantity Top 10

Coupon_Co..



TOP 10 Coupon Code By Quantity With AVG Price

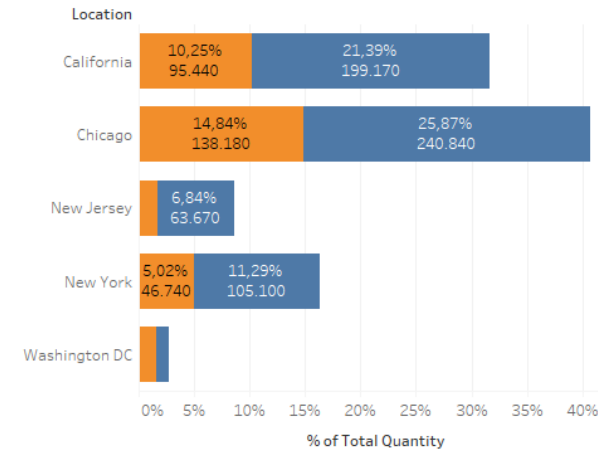


2. Pengalaman Belanja Online Berdasarkan Gender

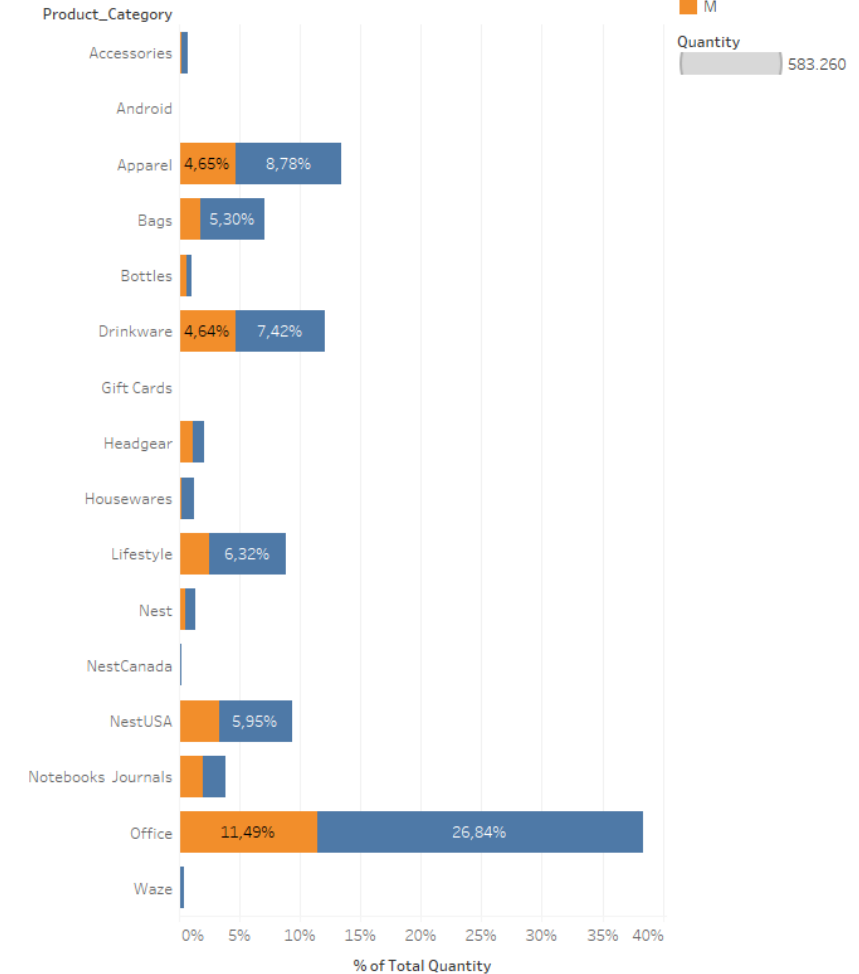
Coupon Code By Quantity With AVG Quantity Top 10



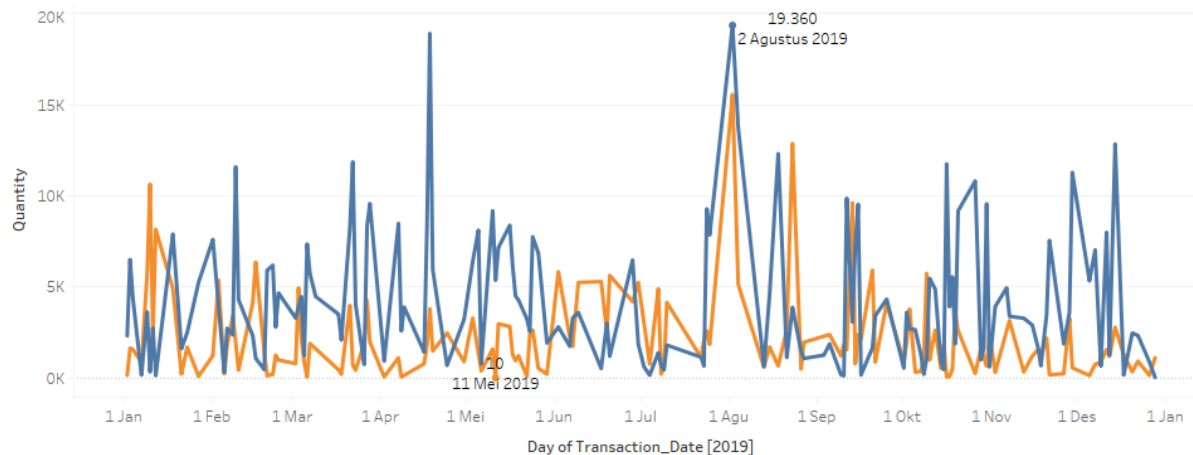
Quantity by location and gender



Market Share by category based gender



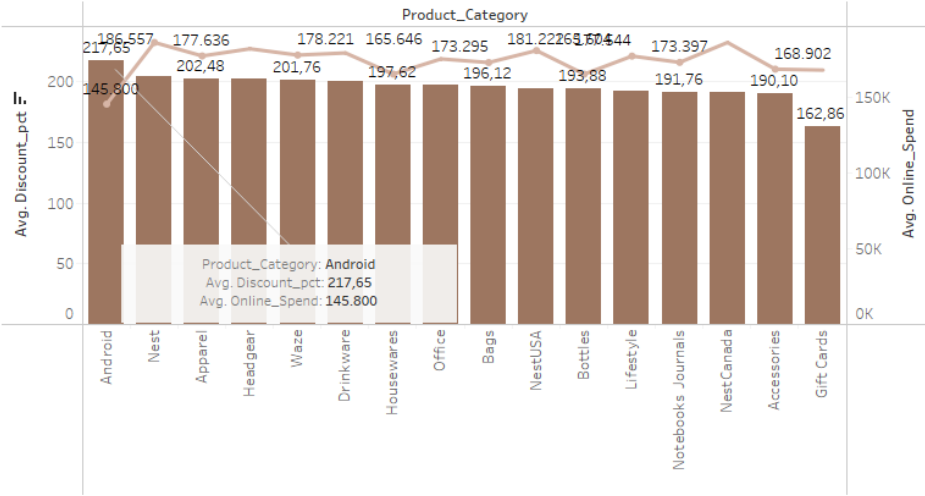
Coupon Code By Quantity With AVG Quantity Top 10



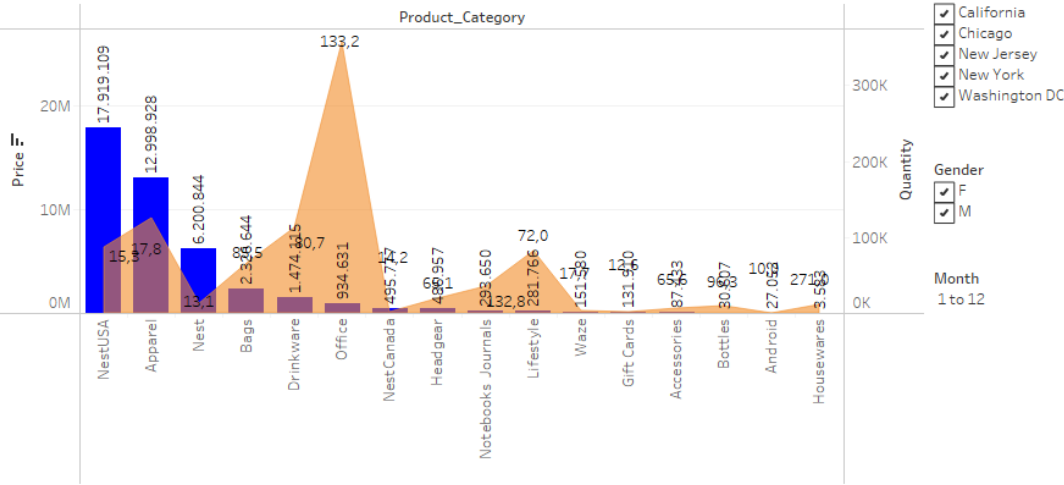


3. Pengaruh Product Category Terhadap Pembelian Customer

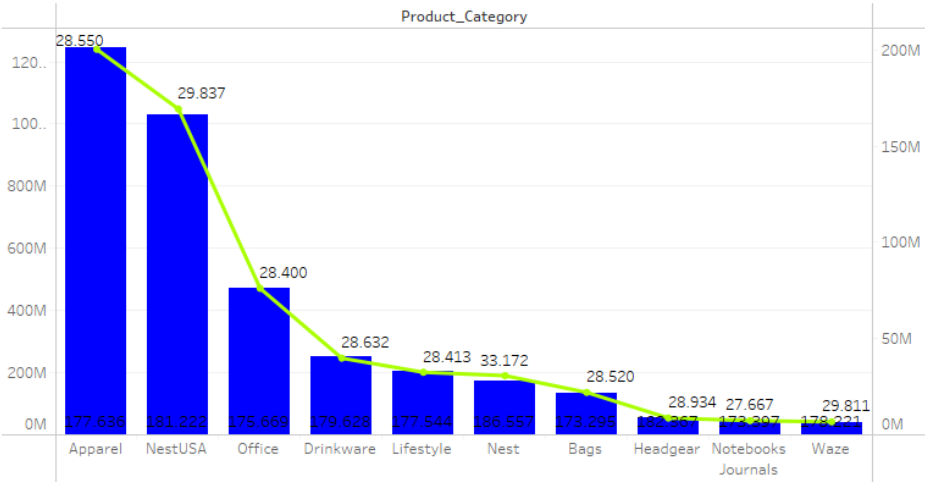
Product Category by AVG Discount_pct with AVG Online_Spend



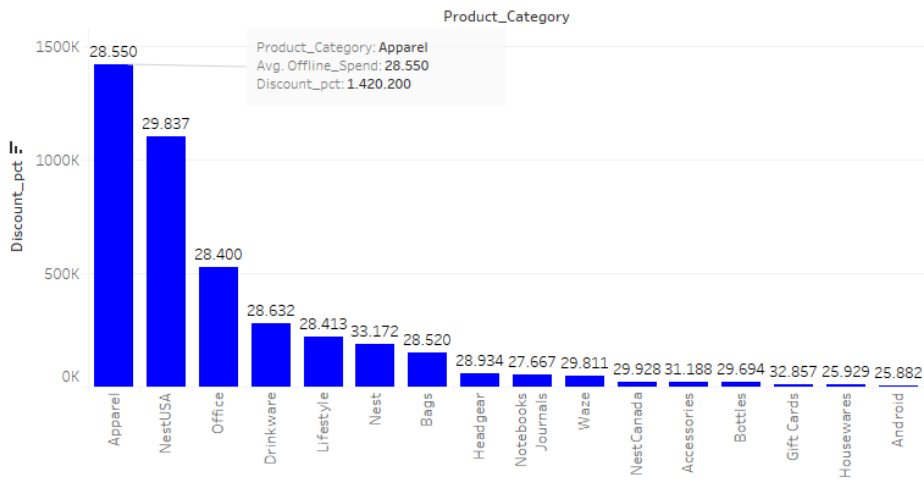
Product Category by Price & AVG Quantity with filters Location All Gender All



Product Category by AVG Online_Spend & AVG Offline_Spend

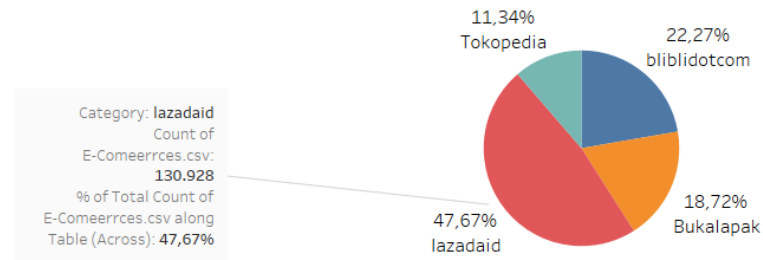


Product Category by Discount_pct with AVG Offline_spend

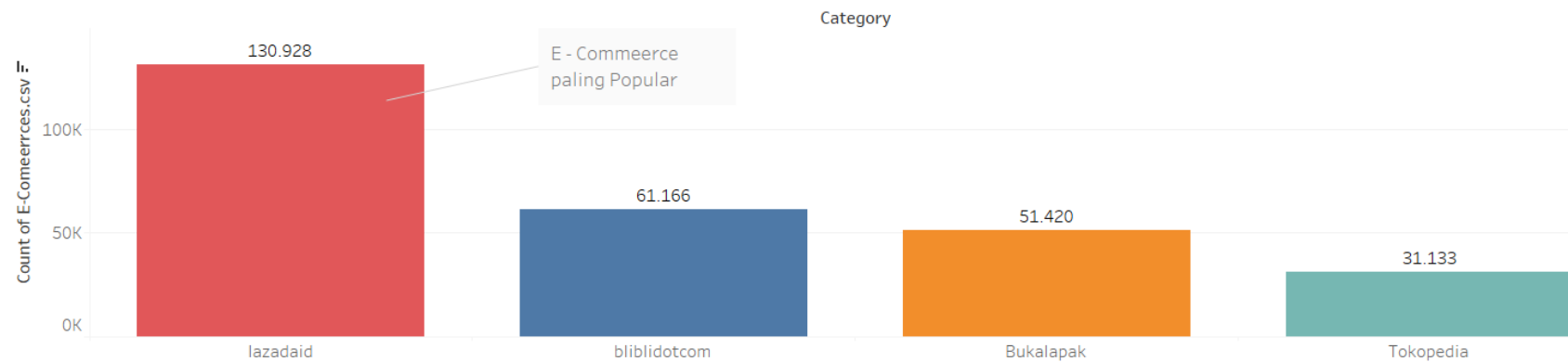


4. Pengaruh Popuplar and Market Share E-Commerce

Market Share E-commerce berdasarkan bulan
January, February, March and 9 more Tahun 2010, 2011, 2012 and 8 more



Popular E-commerce berdasarkan bulan
January, February, March and 9 more Tahun 2010, 2011, 2012 and 8 more



Year of Date

- ☐ Null
- ☒ 2010
- ☒ 2011
- ☒ 2012
- ☒ 2013
- ☒ 2014
- ☒ 2015

Month of Date

- ☐ Null
- ☒ January
- ☒ February
- ☒ March
- ☒ April
- ☒ May
- ☒ June

Category

- ☒ blibliidotcom
- ☒ Bukalapak
- ☒ lazadaid
- ☒ Tokopedia

Count of E-Comeerr..

274.647

Year of Date

- ☐ Null
- ☒ 2010
- ☒ 2011
- ☒ 2012
- ☒ 2013
- ☒ 2014
- ☒ 2015

Month of Date

- ☐ Null
- ☒ January
- ☒ February
- ☒ March
- ☒ April
- ☒ May
- ☒ June



Thank You