



l'école d'ingénierie
informatique

MINI PROJET

**Analyse Des Performances Commerciales :
Insights et Visualisations avec Spark**

DAOUD MOHAMMED-RYAD

SOMMAIRE

TOP selling product	1
Employee Analysis	2
Product delivery duration	3
Orders per mounth	4
Sales analysis	5
Sales Performance By Employee	6
customer life time	7
Products Most Ordered By Vendor	8
Sales Total By Product Category	9
Average Order Value Per Customer	10

NOTE :

L'utilisation de Breeze pour les visualisations dans Scala présente une limitation principale : les graphiques se basent sur des indices plutôt que sur des étiquettes descriptives. Cela rend la lecture et l'interprétation des données, surtout pour les séries temporelles ou catégorielles, moins intuitive. Pour des visualisations plus claires et descriptives, il est recommandé de compléter Breeze avec d'autres outils spécialisés en visualisation de données.

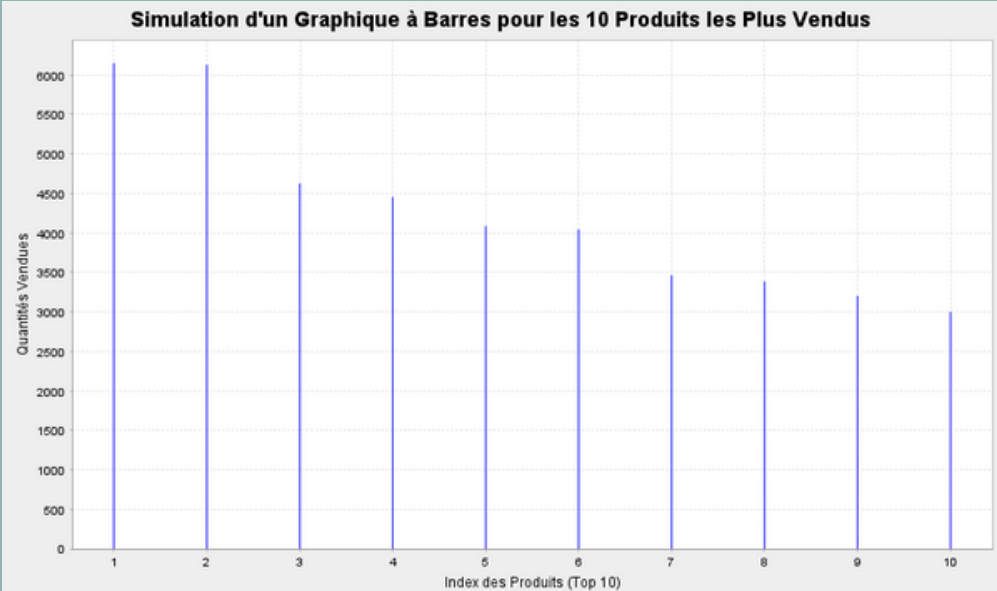
01 - TOP SELLING PRODUCT

l'indicateur recherché est **la quantité totale vendue pour chaque produit**, ce qui permet d'identifier les 10 produits les plus vendus. Cet indicateur fournit une mesure directe de la popularité et de la performance des ventes de chaque produit, en se basant sur la somme des quantités commandées.

AFFICHAGE SPARK

ProductID	ProductName	Total_Quantity_Sold
715	Long-Sleeve Logo ...	6140
712	AWC Logo Cap	6121
711	Sport-100 Helmet,...	4618
708	Sport-100 Helmet,...	4447
864	Classic Vest, S	4079
707	Sport-100 Helmet,...	4036
884	Short-Sleeve Clas...	3455
863	Full-Finger Glove...	3378
714	Long-Sleeve Logo ...	3194
867	Women's Mountain ...	2992

REPRÉSENTATION DE RÉSULTAT



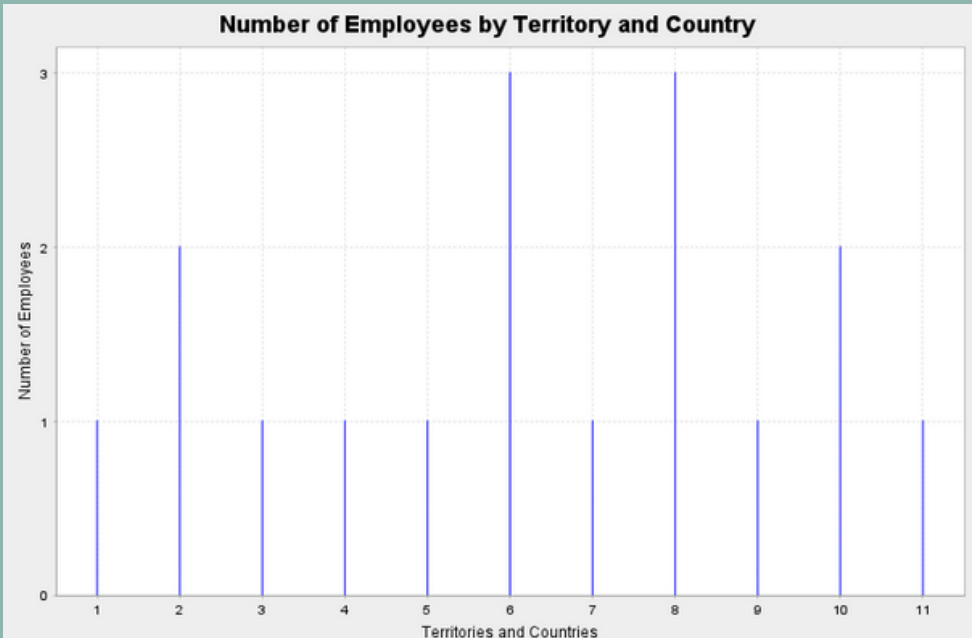
02 - EMPLOYEE ANALYSIS

L'indicateur recherché dans ce code est **le nombre d'employés par région et par pays**. Cet indicateur permet de comprendre la distribution géographique des employés de l'entreprise, offrant une vue d'ensemble de la répartition des effectifs à travers différentes territoires et nations.

AFFICHAGE SPARK

Territory	Country	NumberOfEmployees
Australia	AU	1
Canada	CA	2
Central	US	1
France	FR	1
Germany	DE	1
NULL	NULL	3
Northeast	US	1
Northwest	US	3
Southeast	US	1
Southwest	US	2
United Kingdom	GB	1

REPRÉSENTATION DE RÉSULTAT



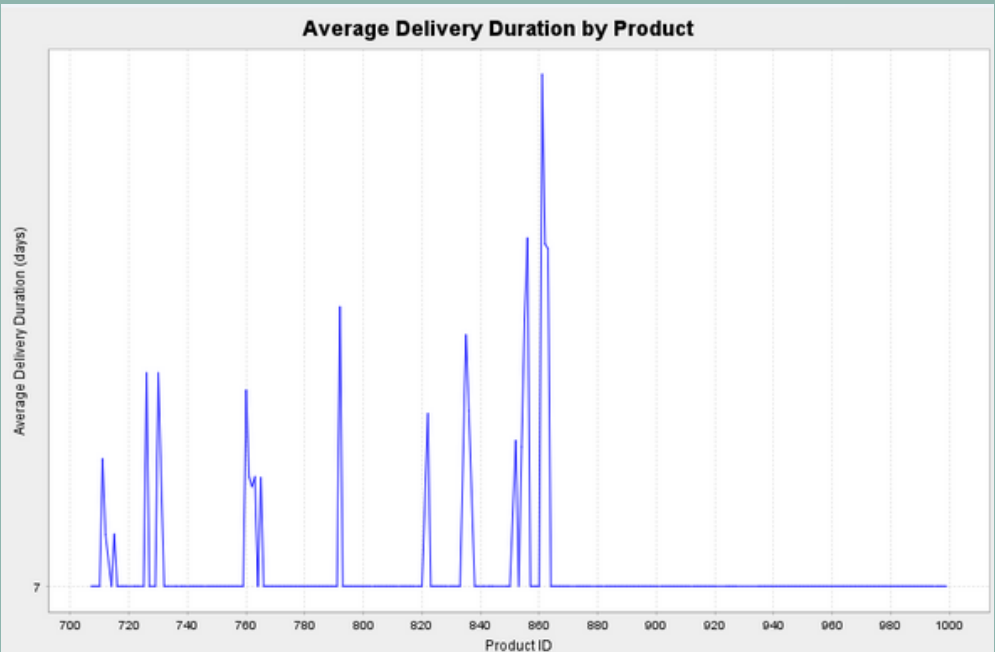
03 - PRODUCT DELIVERY DURATION

l'indicateur recherché est la durée moyenne de livraison pour chaque produit. Ce calcul est effectué en mesurant l'intervalle de temps entre la date de commande (OrderDate) et la date de livraison (ShipDate) pour chaque commande, puis en calculant la moyenne de ces durées pour chaque produit identifié par son ProductID. Cet indicateur est important pour évaluer l'efficacité de la chaîne logistique et la performance de livraison des produits aux clients.

AFFICHAGE SPARK

ProductID	AverageDeliveryDuration
707	7.0
708	7.0
709	7.0
710	7.0
711	7.002072538860103
712	7.000838926174497
714	7.0
715	7.000845308537616
716	7.0
717	7.0
718	7.0
719	7.0
722	7.0
723	7.0
725	7.0

REPRÉSENTATION DE RÉSULTAT



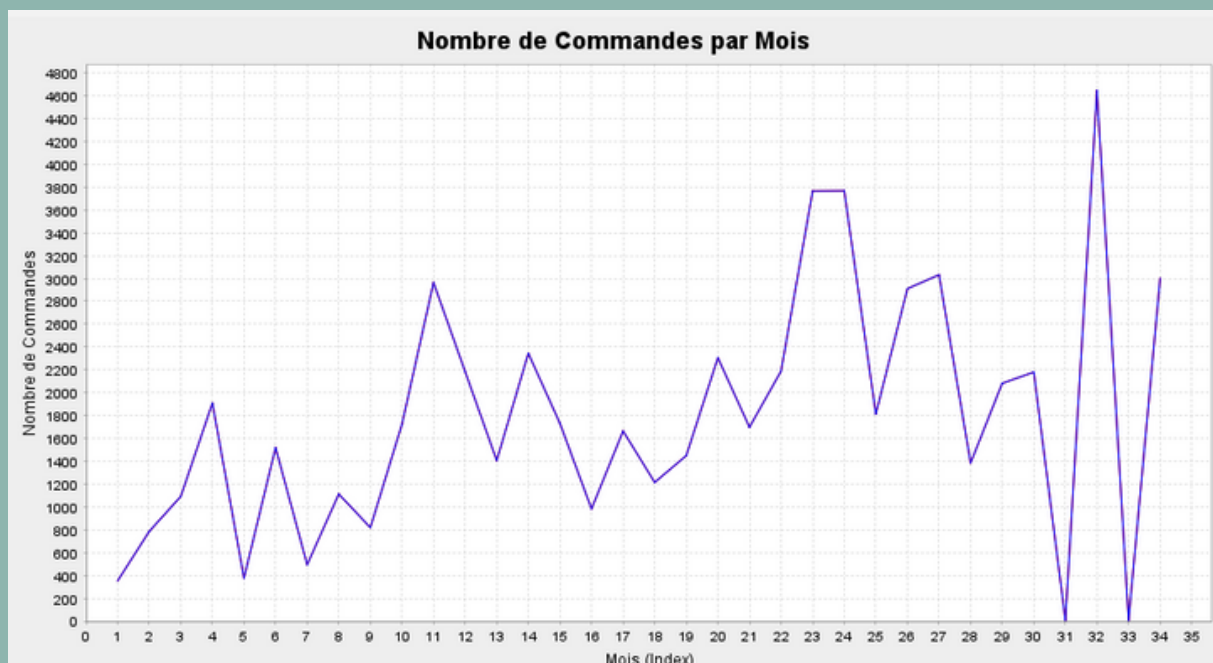
04 - ORDERS PER MOUNTH

L'indicateur calculé dans ce code est **le nombre de commandes par mois et par année**. Cela permet d'évaluer le volume des ventes et la fréquence des commandes sur une base temporelle, offrant une perspective sur la saisonnalité et les tendances de croissance ou de décroissance au fil du temps dans l'activité commerciale.

AFFICHAGE SPARK

Year	Month	NumberOfOrders
2014	5	3002
2014	4	7
2014	3	4647
2014	2	8
2014	1	2180
2013	12	2082
2013	11	1385
2013	10	3032
2013	9	2911
2013	8	1813
2013	7	3768
2013	6	3766
2013	5	2189

REPRÉSENTATION DE RÉSULTAT



05 - SALES ANALYSIS

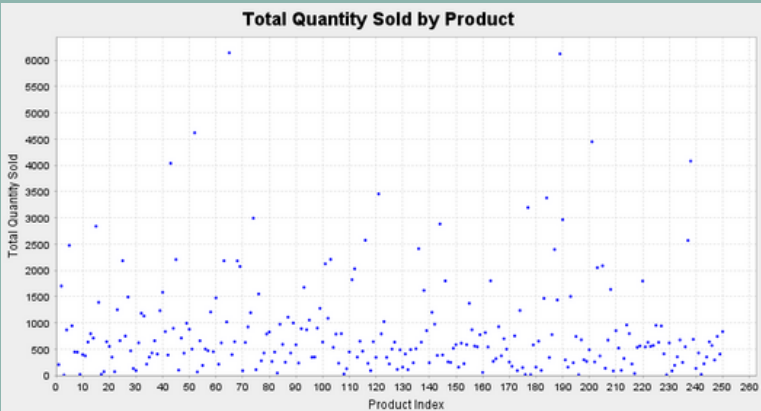
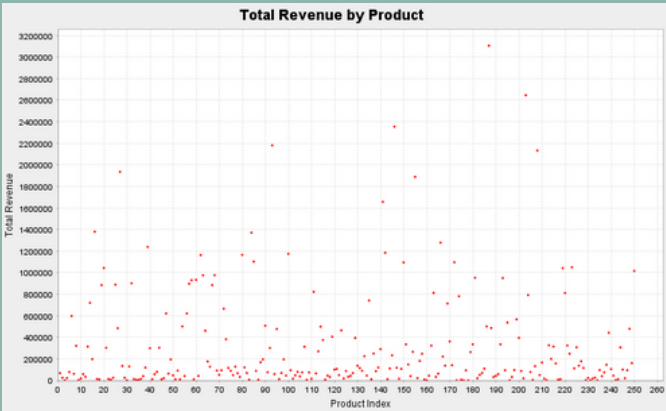
L'indicateur calculé dans ce code est les ventes totales par produit, décomposé en deux mesures spécifiques :

- 1. **Quantité Totale Vendue (Total_Quantity_Sold)** : représente le volume total de chaque produit vendu, offrant un aperçu de la popularité ou de la demande pour chaque produit.
- 2. **Revenu Total (Total_Revenue)** : représente le revenu total généré par les ventes de chaque produit, fournissant une mesure de la contribution financière de chaque produit aux revenus globaux de l'entreprise.

AFFICHAGE SPARK

ProductID	ProductName	Total_Quantity_Sold	Total_Revenue
833	ML Road Frame-W ...	203	68167.51440000003
858	Half-Finger Glove...	1700	24539.430835000014
897	LL Touring Frame ...	4	800.208
808	LL Mountain Handl...	865	22008.268728000006
883	Short-Sleeve Clas...	2474	78280.45733400021
799	Road-550-W Yellow...	943	597453.2714999996
804	HL Fork	444	61034.60962400006
970	Touring-2000 Blue...	441	321027.0281400006
898	LL Touring Frame ...	15	3000.7799999999997
853	Women's Tights, M	394	17727.636
918	LL Mountain Frame...	371	58777.53000000007
857	Men's Bib-Shorts, L	633	34178.20200000001
961	Touring-3000 Yell...	792	314323.2387750005
974	Road-350-W Yellow...	709	720333.7143200005
876	Hitch Rack - 4-Bike	2838	197736.15600000002
976	Road-350-W Yellow...	1390	1380253.8771200012
744	HL Mountain Frame...	17	13765.920000000002

REPRÉSENTATION DE RÉSULTAT



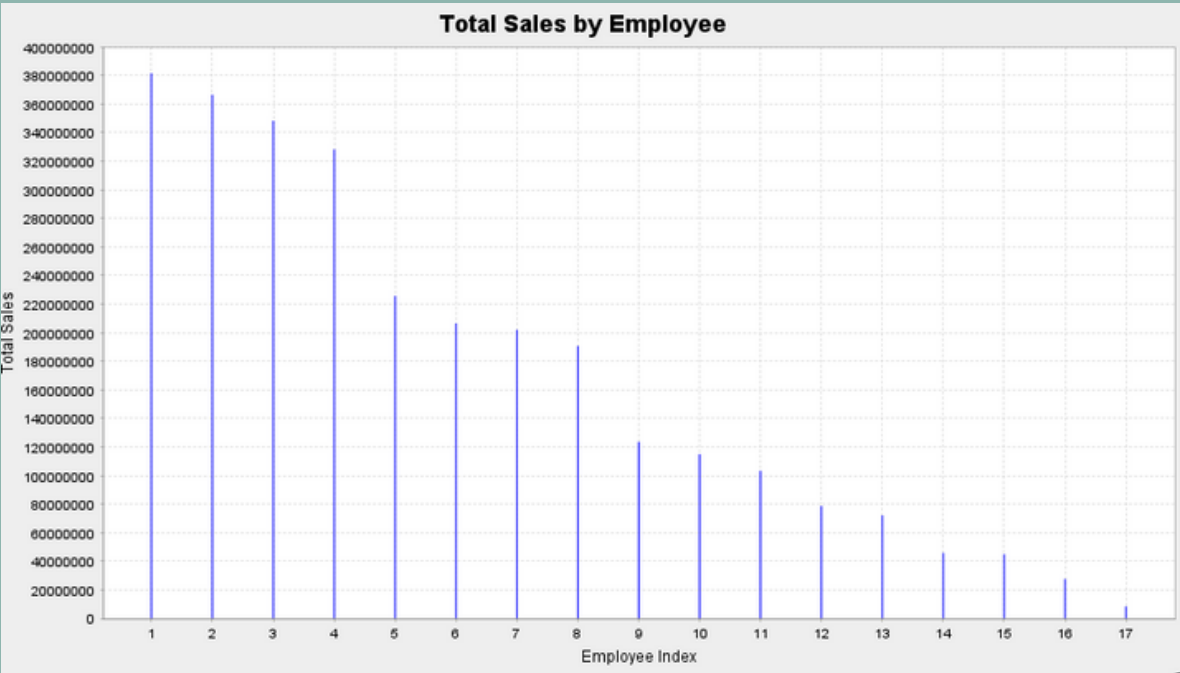
06 - SALES PERFORMANCE BY EMPLOYEE

L'indicateur **montant total des ventes par employé** révèle l'efficacité et la contribution individuelle des employés aux revenus de l'entreprise. En évaluant ce montant, les gestionnaires peuvent identifier les talents de vente clés ainsi que les domaines nécessitant amélioration ou formation. Cela aide à optimiser les stratégies de vente et à encourager une culture de performance au sein de l'équipe. Cet indicateur est essentiel pour la planification des ressources et la reconnaissance des performances exceptionnelles.

AFFICHAGE SPARK

FullName	Total_Sales
Linda Mitchell	3.811255606115012E8
Jae Pak	3.6603733128420067E8
Jillian Carson	3.4786865494949836E8
Michael Blythe	3.278537842759989E8
Shu Ito	2.2525539137440032E8
José Saraiva	2.0604585403839928E8
Tsvi Reiter	2.016128662555986E8
Ranjit Varkey Chu...	1.902185607085009E8
David Campbell	1.2309054527130029E8
Pamela Ansman-Wolfe	1.1451110784209967E8

REPRÉSENTATION DE RÉSULTAT



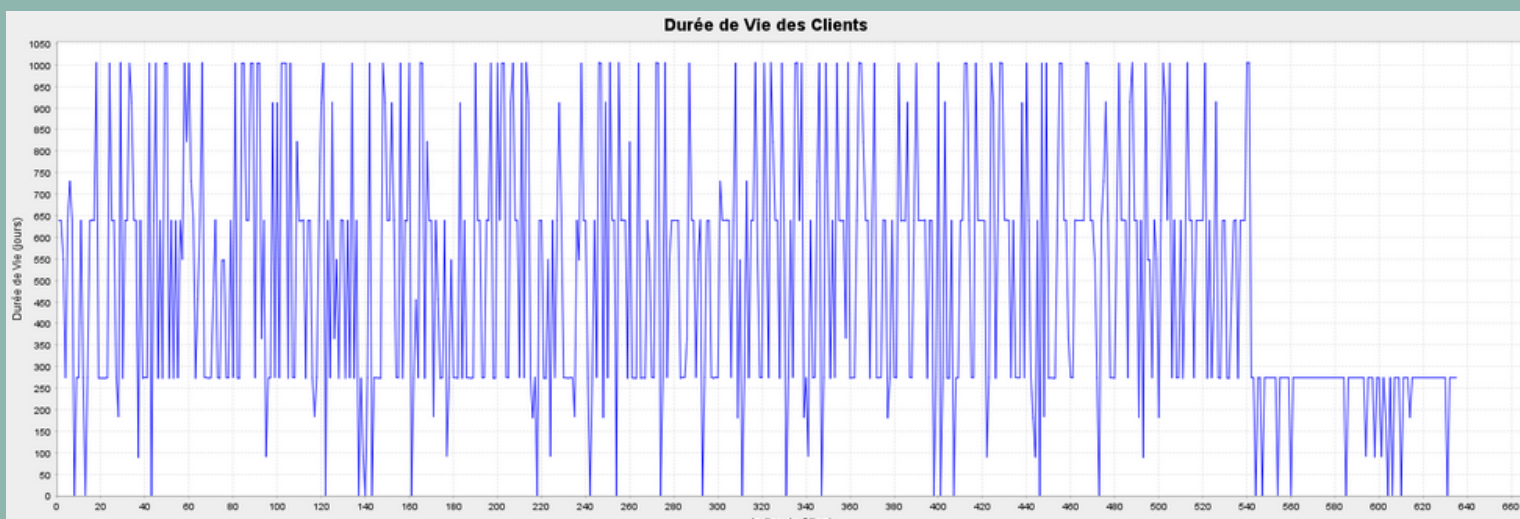
07 - CUSTOMER LIFE TIME

L'indicateur calculé dans ce code est **la durée de vie moyenne d'un client**, mesurée en jours. Ce calcul est effectué en déterminant d'abord la première et la dernière date de commande pour chaque client, puis en calculant l'intervalle entre ces deux dates pour chaque client afin d'obtenir leur durée de vie en tant que client. Finalement, la durée de vie moyenne de tous les clients est calculée et affichée, fournissant une mesure de combien de temps, en moyenne, les clients restent actifs avec l'entreprise.

AFFICHAGE SPARK

CustomerID	DatePremièreCommande	DateDernièreCommande	Durée de Vie Client
1829	2012-07-31	2014-05-01	639
471	2012-07-31	2014-05-01	639
1959	2012-10-30	2014-05-01	548
833	2013-06-30	2014-03-31	274
737	2011-07-01	2013-03-30	638
897	2011-10-01	2013-09-30	730
1025	2012-07-31	2014-05-01	639
1127	2012-08-30	2012-08-30	0
1395	2013-06-30	2014-03-31	274
1483	2013-07-31	2014-05-01	274
1339	2012-07-31	2014-05-01	639
1303	2013-06-30	2014-03-31	274
1903	2013-06-30	2013-06-30	0

REPRÉSENTATION DE RÉSULTAT



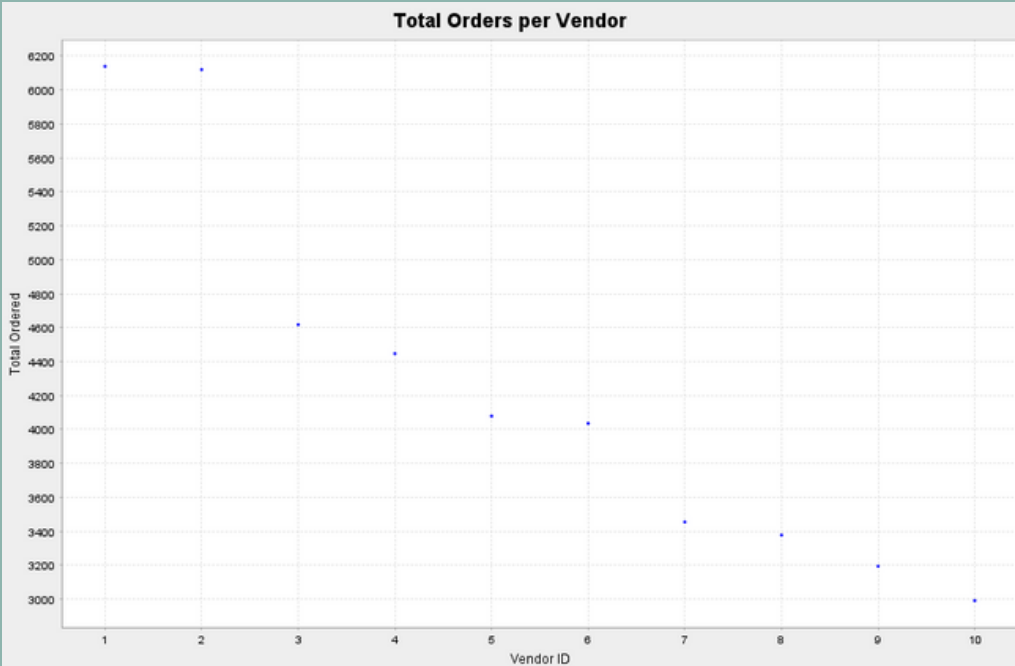
08 - PRODUCTS MOST ORDERED BY VENDOR

l'indicateur calculé est **la quantité totale commandée par produit**, avec une focalisation supplémentaire sur les produits les plus commandés en **les associant à leurs fournisseurs**. Après avoir agrégé les quantités commandées par ProductID, le code joint ces informations avec celles des produits et de leurs fournisseurs pour identifier les 10 produits les plus commandés ainsi que leurs fournisseurs respectifs. Cet indicateur permet non seulement de voir quels produits sont les plus populaires ou les plus demandés, mais aussi de reconnaître les fournisseurs qui contribuent le plus à ces produits à succès.

AFFICHAGE SPARK

VendorID	VendorName	ProductID	TotalOrdered
1636	Integrated Sport ...	715	6140
1636	Integrated Sport ...	712	6121
1520	G & K Bicycle Corp.	711	4618
1520	G & K Bicycle Corp.	708	4447
1676	Team Athletic Co.	864	4079
1520	G & K Bicycle Corp.	707	4036
1636	Integrated Sport ...	884	3455
1676	Team Athletic Co.	863	3378
1636	Integrated Sport ...	714	3194
1676	Team Athletic Co.	867	2992

REPRÉSENTATION DE RÉSULTAT



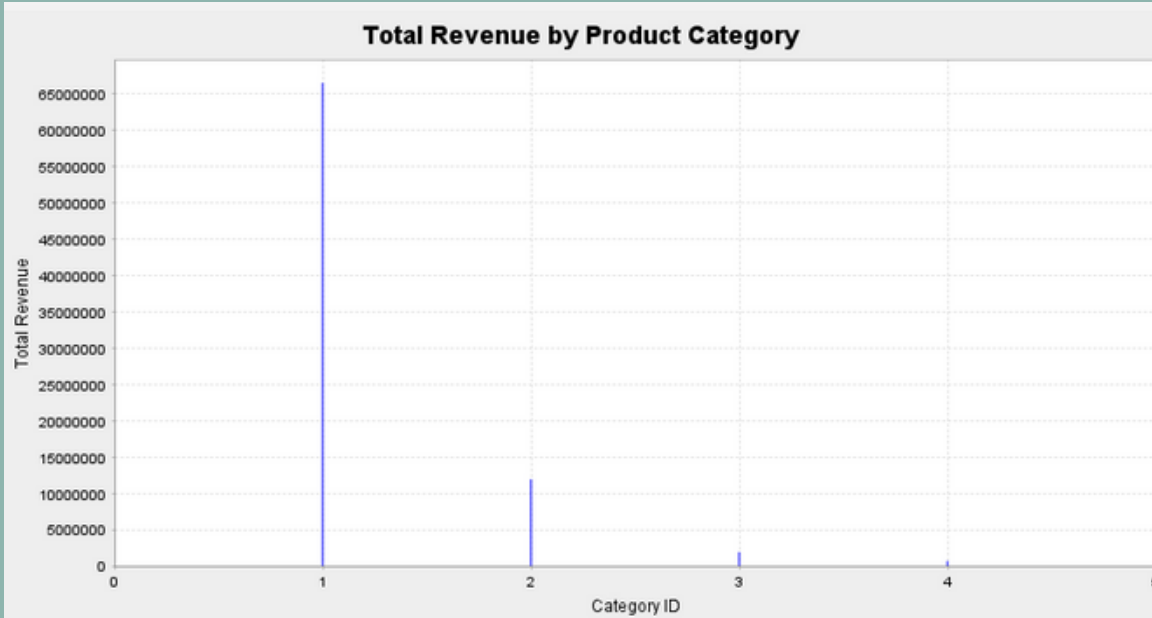
09 - SALES TOTAL BY PRODUCT CATEGORY

l'indicateur calculé est **le revenu total par catégorie de produit**. Le processus comprend plusieurs étapes de jointure de DataFrames pour intégrer des données relatives aux commandes, produits, catégories de produits, et sous-catégories de produits. Finalement, en regroupant les données par CategoryID, le code agrège le total des revenus (LineTotal) pour chaque catégorie de produit. Cela permet d'évaluer la performance financière des différentes catégories de produits de l'entreprise, en identifiant lesquelles génèrent le plus de revenu.

AFFICHAGE SPARK

CategoryID	Total_Revenue
1	6.633302805416833E7
3	1780769.9148009305
4	571312.9239260039
2	1.1802593286434634E7

REPRÉSENTATION DE RÉSULTAT



10 - AVERAGE ORDER VALUE PER CUSTOMER

L'indicateur calculé est **la valeur moyenne des commandes par client**. Ce calcul est réalisé en premier lieu en agrégeant le total des ventes et le nombre de commandes par client, puis en divisant le total des ventes par le nombre de commandes pour chaque client afin de déterminer la valeur moyenne de commande. Ensuite, les résultats sont enrichis par une jointure avec les informations des clients pour inclure le nom complet du client (FullName), permettant d'associer chaque valeur moyenne de commande à un client spécifique. Les résultats sont finalement ordonnés en décroissant selon la valeur moyenne des commandes, mettant en avant les clients avec les valeurs moyennes de commandes les plus élevées.

AFFICHAGE SPARK

CustomerID	FullName	Average_Order_Value
651	Raul Casts	155160.72202758613
1261	Edward Kozłowski	122602.27862827604
615	Joseph Cantoni	121927.10076172852
1125	Phyllis Huntsman	115265.9324517646
1299	Robertson Lee	115102.40251544715
785	Kirk DeGrasse	114516.2802298194
1425	Jane McCarty	104554.17657305957
1403	Frank Mart	102699.47224636377
547	Dave Browning	101118.20402886604
1241	Anton Kirilov	100748.54745714291
625	Carlton Carlisle	99416.674690164
1419	Stacie Mcanich	96722.52022295904
1961	Mandy Vance	96006.40743012371
1735	Karin Lin	84984.77093715730

REPRÉSENTATION DE RÉSULTAT

