

Document de spécification

Vue d'ensemble

- Introduction
- Analyses statistiques de la page
- Statistiques générales
- Analyse par pays
- Analyse par sexe et tranche d'âge
- Analyses statistiques du post
- Recommandations

Introduction

Ce document rassemble la liste des spécifications détaillées recueillies durant vos entretiens avec la community manager.

Veuillez indiquer les résultats de vos requêtes SQL, accompagnés de captures d'écran de vos requêtes SQL.

Dans les pages suivantes :

- les abonnés sont présentés comme : Engaged_Users
- le taux d'engagement est défini de la manière suivante : $\text{Engaged_Fans} + \frac{\text{Engaged_Users}}{\text{somme totale des enregistrements}}$
- la période s'étend du **20 septembre 2018 au 17 octobre 2018**

Statistiques générales

Veuillez inscrire les requêtes SQL pour répondre aux questions suivantes, puis indiquer les résultats dans le tableau ci-dessous :

Question	Votre requête SQL	Résultats
Combien la page a-t-elle enregistré d' utilisateurs engagés au cours de la période donnée ?	<pre>SELECT SUM(Engaged_users) FROM FacebookInsights;</pre>	<div>SUM(Engaged_users)</div> <div>4696375</div>

Quelle est la portée moyenne des posts de la page sur la période donnée ?	<pre>SELECT SUM(Reach) / COUNT(*) AS Portee_moyenne FROM FacebookInsights;</pre>	<div>Portee_moyenne</div> <div>161885</div>
Quel est le taux d'engagement moyen sur la page sur la période donnée ?	<pre>SELECT ROUND((SUM(Engaged_Posts) + SUM(Engaged_Users)) * 1.0 / COUNT(*), 2) AS Taux_engagement_moyen FROM FacebookInsights;</pre>	<div>Taux_engagement_moyen</div> <div>13377.79</div>

Remarques :

- Utilisez le tableau principal FacebookInsights
- Utilisez sum() pour la 1ère requête
- Utilisez sum() / count(*) pour les 2e et 3e requêtes

Statistiques par pays

Question	Votre requête SQL	Résultats																																																																		
Quels sont les 10 premiers codes pays (au vu du nombre maximum de fans) ?	<pre>SELECT CountryCode, MAX(NumberOfFans) as MaxFans FROM FansPerCountry GROUP BY CountryCode ORDER BY MaxFans DESC LIMIT 10;</pre>	<table><thead><tr><th></th><th>CountryCode</th><th>MaxFans</th></tr></thead><tbody><tr><td>1</td><td>CI</td><td>112180</td></tr><tr><td>2</td><td>CM</td><td>102811</td></tr><tr><td>3</td><td>BN</td><td>83561</td></tr><tr><td>4</td><td>FR</td><td>73858</td></tr><tr><td>5</td><td>MG</td><td>72896</td></tr><tr><td>6</td><td>CD</td><td>60706</td></tr><tr><td>7</td><td>BF</td><td>43500</td></tr><tr><td>8</td><td>ML</td><td>40678</td></tr><tr><td>9</td><td>DZ</td><td>39093</td></tr><tr><td>10</td><td>GN</td><td>35521</td></tr></tbody></table>		CountryCode	MaxFans	1	CI	112180	2	CM	102811	3	BN	83561	4	FR	73858	5	MG	72896	6	CD	60706	7	BF	43500	8	ML	40678	9	DZ	39093	10	GN	35521																																	
	CountryCode	MaxFans																																																																		
1	CI	112180																																																																		
2	CM	102811																																																																		
3	BN	83561																																																																		
4	FR	73858																																																																		
5	MG	72896																																																																		
6	CD	60706																																																																		
7	BF	43500																																																																		
8	ML	40678																																																																		
9	DZ	39093																																																																		
10	GN	35521																																																																		
Quels sont les 10 premiers pays (au vu du taux de pénétration : % d'habitants d'un pays qui sont fans) ?	<pre>SELECT FansPerCountry.CountryCode, Population.CountryName, SUM(FansPerCountry.NumberOfFans), Population.Population, ROUND(SUM(FansPerCountry.NumberOfFans) * 1.0 / Population.Population * 100, 2) AS TauxPenetration FROM FansPerCountry JOIN Population ON FansPerCountry.CountryCode = Population.CountryCode GROUP BY FansPerCountry.CountryCode ORDER BY TauxPenetration DESC LIMIT 10;</pre>	<table><thead><tr><th></th><th>CountryCode</th><th>CountryName</th><th>SUM(FansPerCountry.NumberOfFans)</th><th>Population</th><th>TauxPenetration</th></tr></thead><tbody><tr><td>1</td><td>MT</td><td>Mauritius</td><td>69517</td><td>1365300.0</td><td>58.97</td></tr><tr><td>2</td><td>FR</td><td>French Polynesia</td><td>141753</td><td>277879.0</td><td>51.04</td></tr><tr><td>3</td><td>BO</td><td>Boliv Glesion</td><td>137946</td><td>984090.0</td><td>48.56</td></tr><tr><td>4</td><td>GA</td><td>Gabon</td><td>652587</td><td>2119275.0</td><td>30.79</td></tr><tr><td>5</td><td>BM</td><td>Bermuda</td><td>133177</td><td>638388.0</td><td>16.0</td></tr><tr><td>6</td><td>BR</td><td>Brazil</td><td>2294381</td><td>18544360.0</td><td>14.47</td></tr><tr><td>7</td><td>CI</td><td>Cote d'Ivoire</td><td>3166440</td><td>26089280.0</td><td>12.63</td></tr><tr><td>8</td><td>CM</td><td>Cameroon</td><td>2860028</td><td>25216287.0</td><td>11.45</td></tr><tr><td>9</td><td>CG</td><td>Congo, Rep.</td><td>581775</td><td>5644365.0</td><td>10.01</td></tr><tr><td>10</td><td>DZ</td><td>Algeria</td><td>94622</td><td>369820.0</td><td>9.9</td></tr></tbody></table>		CountryCode	CountryName	SUM(FansPerCountry.NumberOfFans)	Population	TauxPenetration	1	MT	Mauritius	69517	1365300.0	58.97	2	FR	French Polynesia	141753	277879.0	51.04	3	BO	Boliv Glesion	137946	984090.0	48.56	4	GA	Gabon	652587	2119275.0	30.79	5	BM	Bermuda	133177	638388.0	16.0	6	BR	Brazil	2294381	18544360.0	14.47	7	CI	Cote d'Ivoire	3166440	26089280.0	12.63	8	CM	Cameroon	2860028	25216287.0	11.45	9	CG	Congo, Rep.	581775	5644365.0	10.01	10	DZ	Algeria	94622	369820.0	9.9
	CountryCode	CountryName	SUM(FansPerCountry.NumberOfFans)	Population	TauxPenetration																																																															
1	MT	Mauritius	69517	1365300.0	58.97																																																															
2	FR	French Polynesia	141753	277879.0	51.04																																																															
3	BO	Boliv Glesion	137946	984090.0	48.56																																																															
4	GA	Gabon	652587	2119275.0	30.79																																																															
5	BM	Bermuda	133177	638388.0	16.0																																																															
6	BR	Brazil	2294381	18544360.0	14.47																																																															
7	CI	Cote d'Ivoire	3166440	26089280.0	12.63																																																															
8	CM	Cameroon	2860028	25216287.0	11.45																																																															
9	CG	Congo, Rep.	581775	5644365.0	10.01																																																															
10	DZ	Algeria	94622	369820.0	9.9																																																															

Statistiques par ville

Question	Votre requête SQL	Résultats																																																							
<p>Quelles sont les 10 villes les moins importantes (au vu du nombre de fans) parmi les pays comprenant plus de 20 millions d'habitants ?</p> <p>Elles peuvent être considérées comme le potentiel de croissance</p>	<pre>SELECT FansPerCity.City, SUM (FansPerCity.NumberOfFans), Population.CountryName, CAST(Population.Population AS INTEGER) AS Population FROM FansPerCity INNER JOIN CityCountry ON FansPerCity.City = CityCountry.City INNER JOIN Population ON CityCountry.CountryCode = Population.CountryCode WHERE Population.Population > 20000000 GROUP BY FansPerCity.City ORDER BY SUM(FansPerCity.NumberOfFans) ASC LIMIT 10;</pre>	<table><tr><th></th><th>City</th><th>SUM (FansPerCity.NumberOfFans)</th><th>CountryName</th><th>Population</th></tr><tr><td>1</td><td>Djéla</td><td>65361</td><td>Algeria</td><td>42828489</td></tr><tr><td>2</td><td>Ngaoundéré</td><td>65364</td><td>Cameroon</td><td>25816237</td></tr><tr><td>3</td><td>Fianarantsoa</td><td>65373</td><td>Madagascar</td><td>25802368</td></tr><tr><td>4</td><td>Tizi Ouzou</td><td>71089</td><td>Algeria</td><td>42828489</td></tr><tr><td>5</td><td>Montréal</td><td>80875</td><td>Canada</td><td>37058866</td></tr><tr><td>6</td><td>Oran</td><td>84004</td><td>Algeria</td><td>42828489</td></tr><tr><td>7</td><td>Bouaké</td><td>96110</td><td>Cote d'Ivoire</td><td>25069269</td></tr><tr><td>8</td><td>Cocody</td><td>103309</td><td>Cote d'Ivoire</td><td>25069269</td></tr><tr><td>9</td><td>Casablanca</td><td>115864</td><td>Morocco</td><td>36082138</td></tr><tr><td>10</td><td>Luanda</td><td>120147</td><td>Angola</td><td>30809762</td></tr></table>		City	SUM (FansPerCity.NumberOfFans)	CountryName	Population	1	Djéla	65361	Algeria	42828489	2	Ngaoundéré	65364	Cameroon	25816237	3	Fianarantsoa	65373	Madagascar	25802368	4	Tizi Ouzou	71089	Algeria	42828489	5	Montréal	80875	Canada	37058866	6	Oran	84004	Algeria	42828489	7	Bouaké	96110	Cote d'Ivoire	25069269	8	Cocody	103309	Cote d'Ivoire	25069269	9	Casablanca	115864	Morocco	36082138	10	Luanda	120147	Angola	30809762
	City	SUM (FansPerCity.NumberOfFans)	CountryName	Population																																																					
1	Djéla	65361	Algeria	42828489																																																					
2	Ngaoundéré	65364	Cameroon	25816237																																																					
3	Fianarantsoa	65373	Madagascar	25802368																																																					
4	Tizi Ouzou	71089	Algeria	42828489																																																					
5	Montréal	80875	Canada	37058866																																																					
6	Oran	84004	Algeria	42828489																																																					
7	Bouaké	96110	Cote d'Ivoire	25069269																																																					
8	Cocody	103309	Cote d'Ivoire	25069269																																																					
9	Casablanca	115864	Morocco	36082138																																																					
10	Luanda	120147	Angola	30809762																																																					

Analyse par sexe et tranche d'âge

Question	Votre requête SQL	Résultats																																																																								
Quelle est la répartition des fans de la page par tranche d'âge (en pourcentage) ?	<pre>SELECT Age, SUM(NumberOfFans), ROUND(SUM(NumberOfFans) * 100.0 / (SELECT SUM(NumberOfFans) FROM FansPerGenderAge), 2) AS Pourcentage FROM FansPerGenderAge GROUP BY Age; ORDER BY Age;</pre>	<table><tr><th>Age</th><th>SUM(NumberOfFans)</th><th>Pourcentage</th></tr><tr><td>13-17</td><td>569718</td><td>2.13</td></tr><tr><td>18-24</td><td>5055032</td><td>21.3</td></tr><tr><td>25-34</td><td>9481862</td><td>38.73</td></tr><tr><td>35-44</td><td>8147038</td><td>32.4</td></tr><tr><td>45-54</td><td>2514170</td><td>9.47</td></tr><tr><td>55-64</td><td>1833118</td><td>5.02</td></tr><tr><td>65+</td><td>1840914</td><td>6.94</td></tr></table>	Age	SUM(NumberOfFans)	Pourcentage	13-17	569718	2.13	18-24	5055032	21.3	25-34	9481862	38.73	35-44	8147038	32.4	45-54	2514170	9.47	55-64	1833118	5.02	65+	1840914	6.94																																																
Age	SUM(NumberOfFans)	Pourcentage																																																																								
13-17	569718	2.13																																																																								
18-24	5055032	21.3																																																																								
25-34	9481862	38.73																																																																								
35-44	8147038	32.4																																																																								
45-54	2514170	9.47																																																																								
55-64	1833118	5.02																																																																								
65+	1840914	6.94																																																																								
Quelle est la répartition des fans de la page par sexe (en pourcentage) ?	<pre>SELECT Gender, SUM(NumberOfFans), ROUND(SUM(NumberOfFans) * 100.0 / (SELECT SUM(NumberOfFans) FROM FansPerGenderAge), 2) AS Pourcentage FROM FansPerGenderAge GROUP BY Gender;</pre>	<table><tr><th>Gender</th><th>SUM(NumberOfFans)</th><th>Pourcentage</th></tr><tr><td>F</td><td>14981907</td><td>56.46</td></tr><tr><td>M</td><td>11527839</td><td>43.44</td></tr><tr><td>U</td><td>25126</td><td>0.09</td></tr></table>	Gender	SUM(NumberOfFans)	Pourcentage	F	14981907	56.46	M	11527839	43.44	U	25126	0.09																																																												
Gender	SUM(NumberOfFans)	Pourcentage																																																																								
F	14981907	56.46																																																																								
M	11527839	43.44																																																																								
U	25126	0.09																																																																								
Disposons-nous du même nombre de jours pour tous les sexes et tranches d'âge ?	<pre>SELECT Gender, Age, COUNT(DISTINCT Date) AS nombre_de_jours FROM FansPerGenderAge GROUP BY Gender, Age ORDER BY nombre_de_jours;</pre>	<table><tr><th>Gender</th><th>Age</th><th>NombreDeJours</th><th>M</th><th>13-17</th><th>28</th><th>U</th><th>13-17</th><th>28</th></tr><tr><td>F</td><td>13-17</td><td>28</td><td>M</td><td>18-24</td><td>28</td><td>U</td><td>18-24</td><td>28</td></tr><tr><td>F</td><td>18-24</td><td>28</td><td>M</td><td>25-34</td><td>28</td><td>U</td><td>25-34</td><td>28</td></tr><tr><td>F</td><td>25-34</td><td>28</td><td>M</td><td>35-44</td><td>28</td><td>U</td><td>35-44</td><td>28</td></tr><tr><td>F</td><td>35-44</td><td>28</td><td>M</td><td>45-54</td><td>28</td><td>U</td><td>45-54</td><td>28</td></tr><tr><td>F</td><td>45-54</td><td>28</td><td>M</td><td>55-64</td><td>28</td><td>U</td><td>55-64</td><td>28</td></tr><tr><td>F</td><td>55-64</td><td>28</td><td>M</td><td>65+</td><td>28</td><td>U</td><td>65+</td><td>28</td></tr><tr><td>F</td><td>65+</td><td>28</td><td>M</td><td>65+</td><td>28</td><td>U</td><td>65+</td><td>28</td></tr></table>	Gender	Age	NombreDeJours	M	13-17	28	U	13-17	28	F	13-17	28	M	18-24	28	U	18-24	28	F	18-24	28	M	25-34	28	U	25-34	28	F	25-34	28	M	35-44	28	U	35-44	28	F	35-44	28	M	45-54	28	U	45-54	28	F	45-54	28	M	55-64	28	U	55-64	28	F	55-64	28	M	65+	28	U	65+	28	F	65+	28	M	65+	28	U	65+	28
Gender	Age	NombreDeJours	M	13-17	28	U	13-17	28																																																																		
F	13-17	28	M	18-24	28	U	18-24	28																																																																		
F	18-24	28	M	25-34	28	U	25-34	28																																																																		
F	25-34	28	M	35-44	28	U	35-44	28																																																																		
F	35-44	28	M	45-54	28	U	45-54	28																																																																		
F	45-54	28	M	55-64	28	U	55-64	28																																																																		
F	55-64	28	M	65+	28	U	65+	28																																																																		
F	65+	28	M	65+	28	U	65+	28																																																																		

Remarques

Vous devez diviser le champ SexeAge en 2 nouvelles colonnes

Créez d'abord les 2 nouvelles colonnes

- Alter table FansPerGenderAge add Gender char(1);
- Alter table FansPerGenderAge add Age varchar(8);

Vérifiez que les colonnes ont été créées avec

- pragma table_info('FansPerGenderAge');

Puis remplissez ces colonnes avec les bonnes analyses de la colonne SexeAge

- Update FansPerGenderAge set Gender = substr(SexeAge, 1,1);
- Update FansPerGenderAge set Age = substr(SexeAge, 3,5);

Confirmez avec

- select * from FansPerGenderAge limit 2;

Engagement dans le temps

Question	Votre requête SQL	Résultats																																								
Quelle est la répartition du taux d'engagement des posts selon les jours de la semaine (lundi, mardi,...) ?	<pre>SELECT jour_semaine, strftime('%w', Date) AS Jour, ROUND(CAST(SUM(Engaged_Users + Engaged_Fans) AS REAL) / (SELECT CAST(SUM(Engaged_Users + Engaged_Fans) AS REAL) FROM FacebookInsights), 2) AS Taux_engagement FROM FacebookInsights GROUP BY strftime('%w', Date) ORDER BY Jour;</pre>	<table><tr><th>jour_semaine</th><th>Jour</th><th>Taux_engagement</th></tr><tr><td>Dimanche</td><td>0</td><td>0.12</td></tr><tr><td>Lundi</td><td>1</td><td>0.15</td></tr><tr><td>Mardi</td><td>2</td><td>0.15</td></tr><tr><td>Mercredi</td><td>3</td><td>0.22</td></tr><tr><td>Jeudi</td><td>4</td><td>0.11</td></tr><tr><td>Vendredi</td><td>5</td><td>0.13</td></tr><tr><td>Samedi</td><td>6</td><td>0.12</td></tr></table>	jour_semaine	Jour	Taux_engagement	Dimanche	0	0.12	Lundi	1	0.15	Mardi	2	0.15	Mercredi	3	0.22	Jeudi	4	0.11	Vendredi	5	0.13	Samedi	6	0.12																
jour_semaine	Jour	Taux_engagement																																								
Dimanche	0	0.12																																								
Lundi	1	0.15																																								
Mardi	2	0.15																																								
Mercredi	3	0.22																																								
Jeudi	4	0.11																																								
Vendredi	5	0.13																																								
Samedi	6	0.12																																								
Quel est le meilleur jour de la semaine pour publier des posts ?	<pre>SELECT jour_semaine, strftime('%w', Date) AS Jour, ROUND(CAST(SUM(Engaged_Users + Engaged_Fans) AS REAL) / (SELECT CAST(SUM(Engaged_Users + Engaged_Fans) AS REAL) FROM FacebookInsights), 2) AS Taux_engagement FROM FacebookInsights GROUP BY strftime('%w', Date) ORDER BY Taux_engagement DESC LIMIT 1;</pre>	<table><tr><th>jour_semaine</th><th>Jour</th><th>Taux_engagement</th></tr><tr><td>Mercredi</td><td>3</td><td>0.22</td></tr></table>	jour_semaine	Jour	Taux_engagement	Mercredi	3	0.22																																		
jour_semaine	Jour	Taux_engagement																																								
Mercredi	3	0.22																																								
Quelle est la répartition du taux d'engagement des posts selon les heures de la journée ?	<pre>SELECT tranche_horaire, ROUND(SUM(Engaged_Fans + Engaged_Users) * 100.0 / (SELECT SUM(Engaged_Fans + Engaged_Users) FROM FacebookInsights), 2) AS taux_engagement_pourcent FROM FacebookInsights GROUP BY tranche_horaire ORDER BY taux_engagement_pourcent DESC;</pre>	<table><tr><th>tranche_horaire</th><th>taux_engagement_pourcent</th></tr><tr><td>06:00-06:00</td><td>18.68</td></tr><tr><td>06:00-07:00</td><td>6.68</td></tr><tr><td>11:00-12:00</td><td>6.11</td></tr><tr><td>07:00-08:00</td><td>6.9</td></tr><tr><td>08:00-09:00</td><td>5.89</td></tr><tr><td>09:00-10:00</td><td>5.64</td></tr><tr><td>12:00-13:00</td><td>6.49</td></tr><tr><td>17:00-18:00</td><td>0.37</td></tr><tr><td>18:00-19:00</td><td>4.99</td></tr><tr><td>19:00-20:00</td><td>4.87</td></tr><tr><td>16:00-17:00</td><td>4.5</td></tr><tr><td>20:00-21:00</td><td>4.45</td></tr><tr><td>20:00-21:00</td><td>4.39</td></tr><tr><td>10:00-11:00</td><td>4.36</td></tr><tr><td>18:00-19:00</td><td>4.08</td></tr><tr><td>21:00-22:00</td><td>3.88</td></tr><tr><td>14:00-15:00</td><td>3.76</td></tr><tr><td>22:00-23:00</td><td>3.5</td></tr><tr><td>16:00-17:00</td><td>2.17</td></tr></table>	tranche_horaire	taux_engagement_pourcent	06:00-06:00	18.68	06:00-07:00	6.68	11:00-12:00	6.11	07:00-08:00	6.9	08:00-09:00	5.89	09:00-10:00	5.64	12:00-13:00	6.49	17:00-18:00	0.37	18:00-19:00	4.99	19:00-20:00	4.87	16:00-17:00	4.5	20:00-21:00	4.45	20:00-21:00	4.39	10:00-11:00	4.36	18:00-19:00	4.08	21:00-22:00	3.88	14:00-15:00	3.76	22:00-23:00	3.5	16:00-17:00	2.17
tranche_horaire	taux_engagement_pourcent																																									
06:00-06:00	18.68																																									
06:00-07:00	6.68																																									
11:00-12:00	6.11																																									
07:00-08:00	6.9																																									
08:00-09:00	5.89																																									
09:00-10:00	5.64																																									
12:00-13:00	6.49																																									
17:00-18:00	0.37																																									
18:00-19:00	4.99																																									
19:00-20:00	4.87																																									
16:00-17:00	4.5																																									
20:00-21:00	4.45																																									
20:00-21:00	4.39																																									
10:00-11:00	4.36																																									
18:00-19:00	4.08																																									
21:00-22:00	3.88																																									
14:00-15:00	3.76																																									
22:00-23:00	3.5																																									
16:00-17:00	2.17																																									

Quelle est la meilleure heure de la journée pour publier des posts ?

```
SELECT
  tranche_horaire,
  ROUND(
    SUM(Engaged_Posts * Engaged_Users) * 100.0 /
    (SELECT SUM(Engaged_Posts * Engaged_Users) FROM FacebookInsights),
    2) AS taux_engagement_pourcent
FROM FacebookInsights
GROUP BY tranche_horaire
ORDER BY taux_engagement_pourcent DESC
LIMIT 1;
```

tranche_horaire	taux_engagement_pourcent
08:00-08:00	18.62

Remarques:

- Vous travaillez sur le tableau FacebookInsights. La colonne created_time n'est pas dans le bon format et ne peut pas être analysée avec strftime.
- Vous devez d'abord créer une colonne Date en analysant la colonne created_time. Cela implique la concaténation de chaînes avec ||
 - Alter table FacebookInsights add Date varchar(24);
 - Update FacebookInsights set Date = substr(created_time, 7,4) || '-' || substr(created_time, 4,2) || '-' || substr(created_time, 1,2) ;
- La fonction strftime est très utile pour analyser la date : https://www.sqlite.org/lang_datefunc.html

Recommandations

- Première recommandation
- Deuxième recommandation
- Recommandation supplémentaire (facultative)