



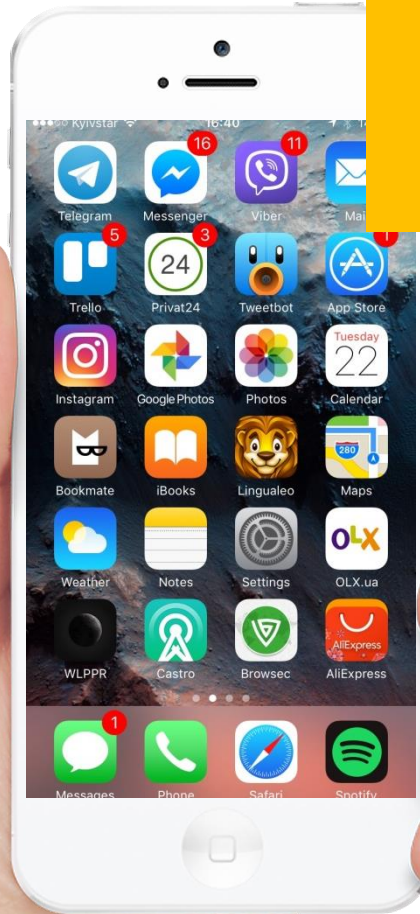
# Analysis of Google Play Store apps

---

---

By Alina Hohryakova

Evaluation parameters are app's average number of reactions (count of rating scores assigned) and / or number installs, and correlation between rating and the percentage of rating count to installs.



# Relevance of the topic



Which applications in today's market have the most relevance perspective;



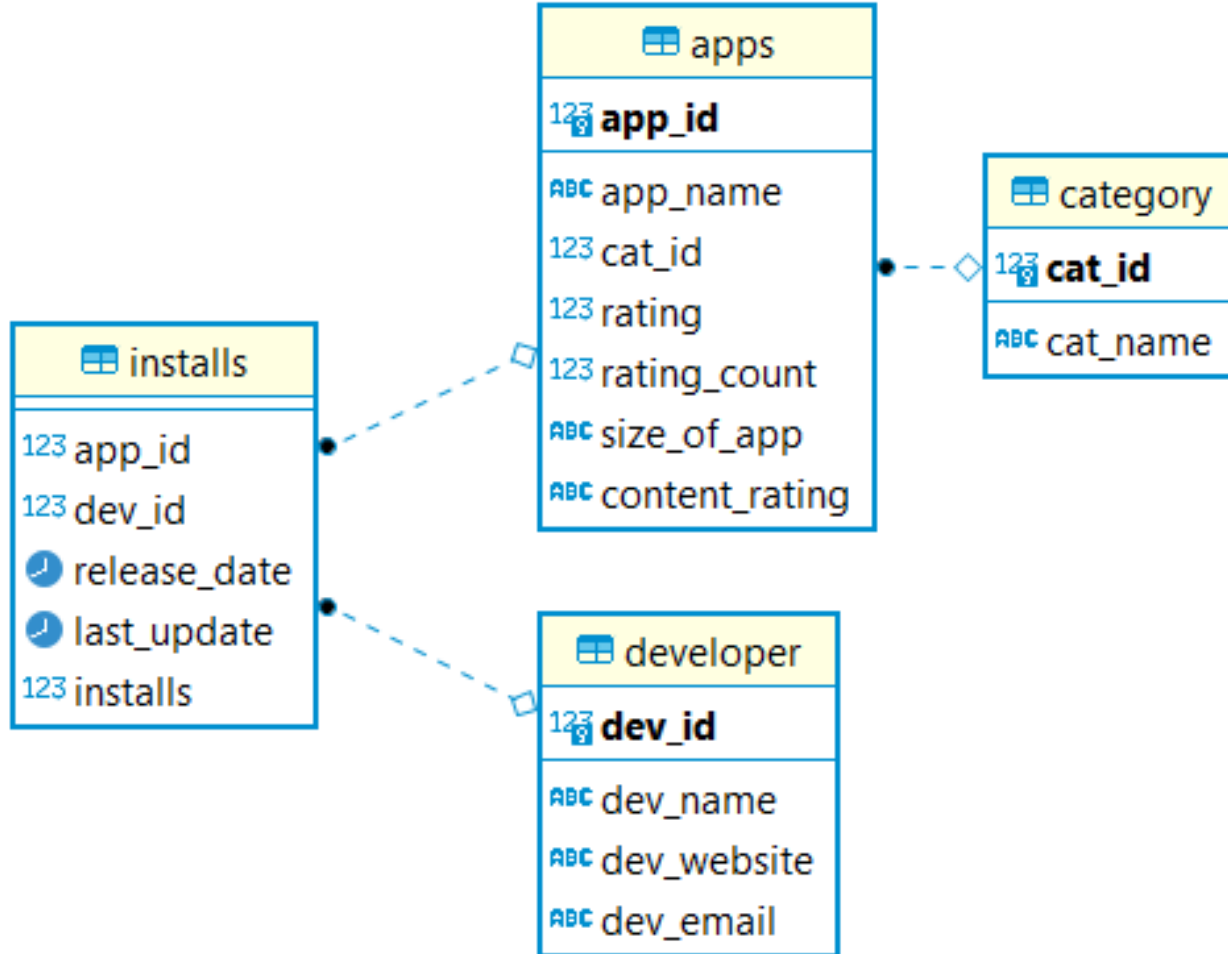
Which apps characteristics are better perceived by the market.

# Data sources

---

kaggle





# DATABASE SCHEMA

SNOWFLAKE

4 TABLES

1 FACT TABLE

3 DIMENSION TABLES



# Popular apps with bad realization

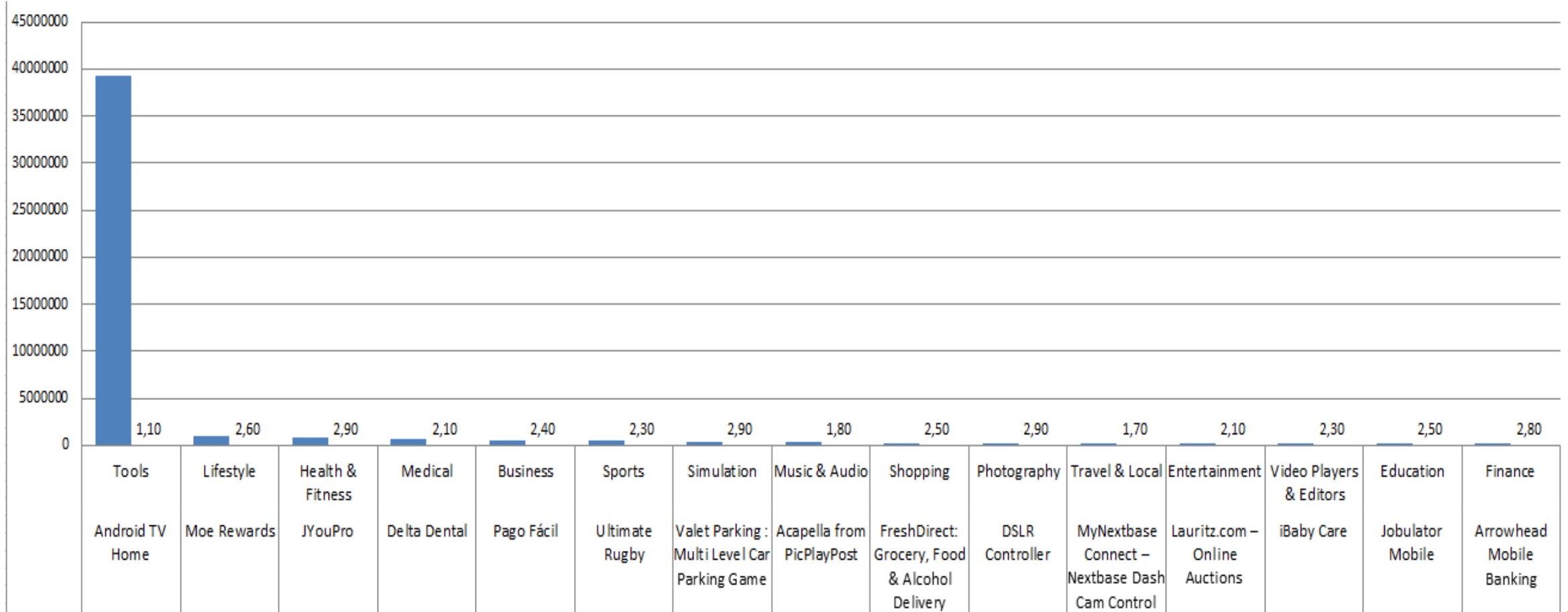
Surprisingly, the most downloaded app came out with the lowest rating, and this app is from the category of tools. Thus, we learned about the prospect of developing a quality application from the category of tools.

```
WITH cte as
  (SELECT a.app_name ,c.cat_id ,c.cat_name , min(a.rating) AS rating , i.installs
   FROM installs i
   INNER JOIN apps a
   ON i.app_id =a.app_id
   INNER JOIN category c
   ON a.cat_id =c.cat_id
   WHERE a.rating < 3 AND i.installs > 100000
   GROUP BY a.app_name ,c.cat_id ,c.cat_name ,i.installs
   ORDER BY c.cat_name )

SELECT c.app_name, c.cat_name, c.installs,c.rating
FROM cte c
LEFT JOIN cte t
ON c.cat_id=t.cat_id
AND c.installs<t.installs
WHERE t.cat_id IS NULL
ORDER BY installs DESC,cat_name;
```

	ABC app_name	ABC cat_name	123 installs	123 rating
1	Android TV Home	Tools	39 240 688	1,1
2	Moe Rewards	Lifestyle	991 584	2,6
3	JYouPro	Health & Fitness	860 488	2,9
4	Delta Dental	Medical	600 291	2,1
5	Pago Fácil	Business	586 558	2,4
6	Ultimate Rugby	Sports	536 055	2,3
7	Valet Parking : Multi	Simulation	403 130	2,9
8	Acapella from PicPla	Music & Audio	357 543	1,8
9	FreshDirect: Grocery,	Shopping	191 388	2,5
10	DSLR Controller	Photography	180 223	2,9
11	MyNextbase Connect	Travel & Local	175 833	1,7
12	Lauritz.com – Online	Entertainment	155 166	2,1
13	iBaby Care	Video Players & E	143 051	2,3
14	Jobulator Mobile	Education	139 566	2,5
15	Arrowhead Mobile B	Finance	112 268	2,8

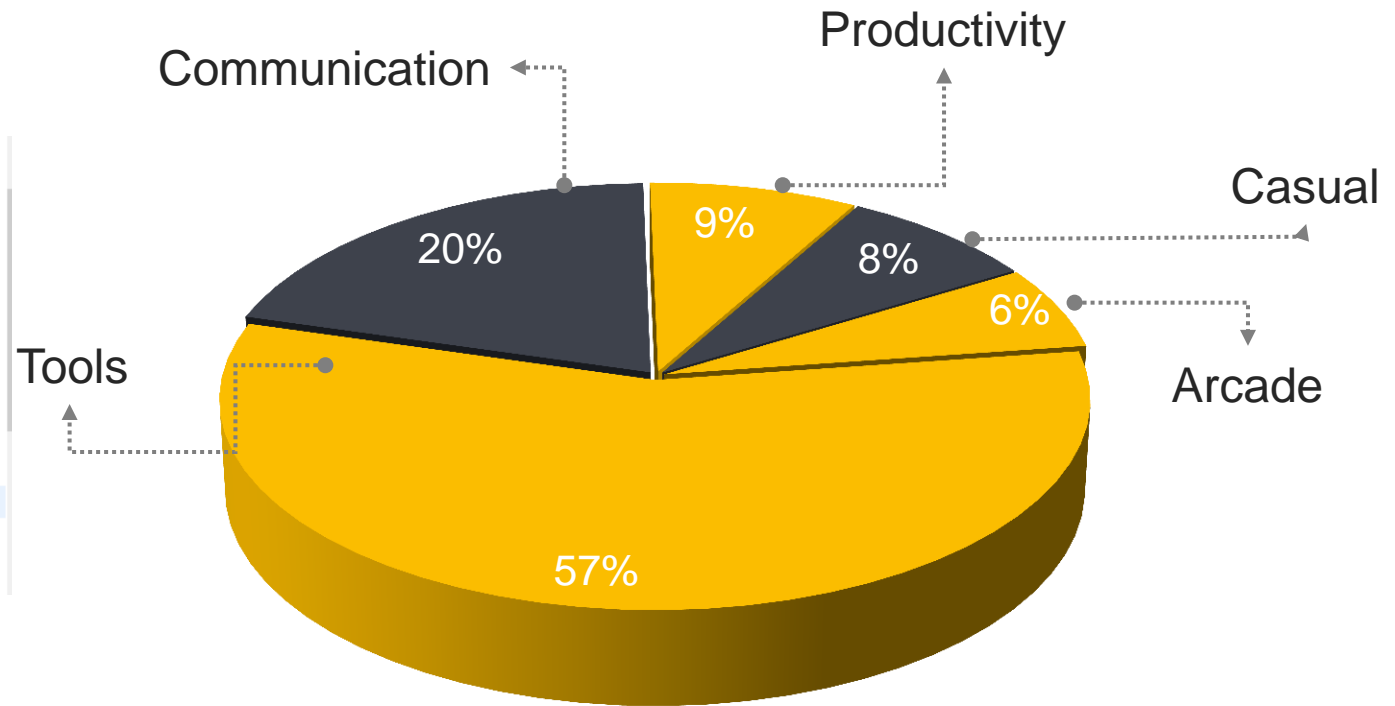
# Popular apps with bad realization



# 5 the most popular apps category by the amount of its installed

```
SELECT category, sum_of_installs
FROM
(SELECT c.cat_name AS category ,sum(i.installs) AS sum_of_installs,
RANK() OVER(ORDER BY sum(i.installs) DESC) AS rank_of_category
FROM category c
INNER JOIN apps a
ON c.cat_id=a.cat_id
INNER JOIN installs i
ON a.app_id =i.app_id
GROUP BY category
ORDER BY rank_of_category,c.cat_name
) qwe
WHERE rank_of_category < 6
ORDER BY sum_of_installs DESC,category
```

	category	sum_of_installs
1	Tools	9 788 157 082
2	Communication	3 492 611 760
3	Productivity	1 468 517 085
4	Casual	1 412 009 416
5	Arcade	1 108 301 385



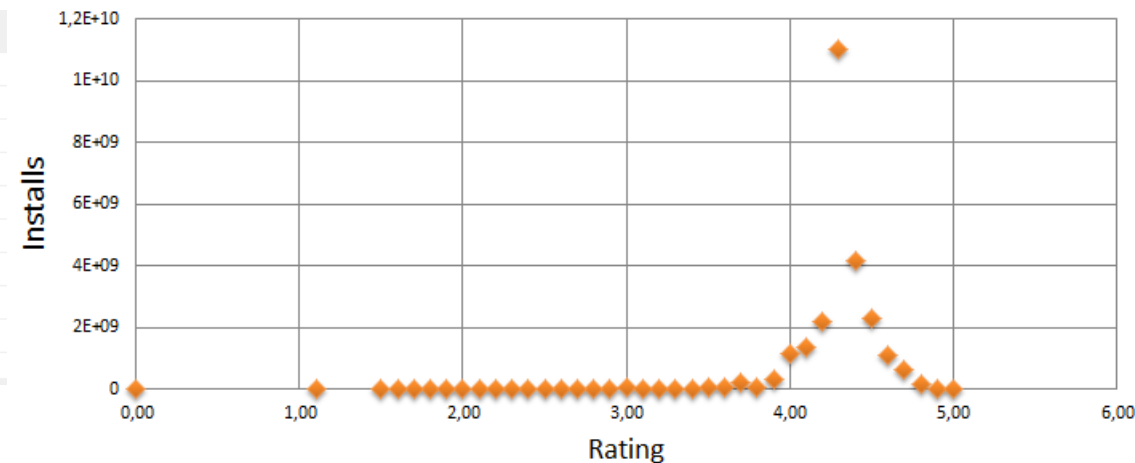
# The correlation between rating and installs

In this query we can see that the most installed apps have pretty high rating.

```
SELECT rating, sum(installs) AS sum_of_installs
FROM
  (SELECT a.rating ,i.installs
   FROM apps a
   INNER JOIN installs i
   ON a.app_id =i.app_id
  ) qwe
GROUP BY rating
ORDER BY sum_of_installs DESC
```

	rating	sum_of_installs
1	4,3	11 007 928 489
2	4,4	4 160 415 401
3	4,5	2 297 042 549
4	4,2	2 202 559 478
5	4,1	1 377 403 001
6	4	1 150 108 909
7	4,6	1 108 741 854
8	4,7	646 182 526
9	3,9	318 600 552
10	3,7	238 053 376

the correlation between rating and installs





# The correlation between rating and the percentage of rating count to installs

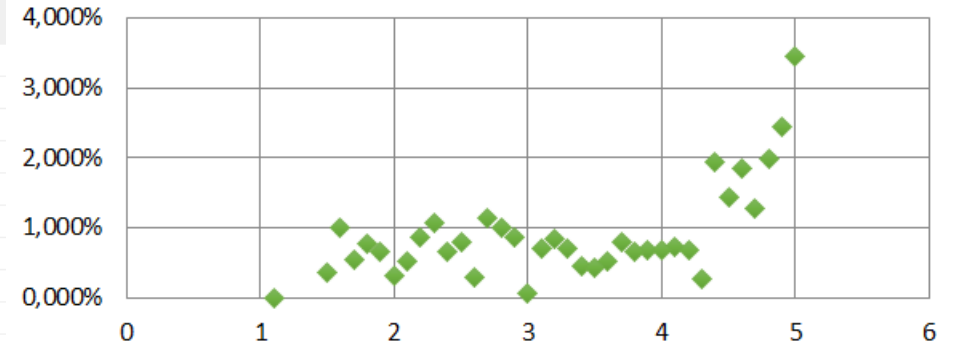
In conclusion, we can note that the highest rated apps have the highest percentage of feedback, but the number of people who have rated the app is still very small relative to the number of installs.

```
-- 4 correlation between rating and the percentage of rating count to installs
SELECT rating, concat(to_char((SUM(rating_count)*100 / sum(sum(installs))
                        OVER (PARTITION BY rating)), '99.999'), ' %') as percentage
FROM
  (SELECT a.rating ,a.rating_count, i.installs
   FROM apps a
   INNER JOIN installs i
   ON a.app_id =i.app_id

   ) qwe
WHERE rating > 0
GROUP BY rating
ORDER BY percentage DESC
```

rating	percentage
5	3.442 %
4,9	2.445 %
4,8	1.975 %
4,4	1.949 %
4,6	1.843 %
4,5	1.435 %
4,7	1.280 %
2,7	1.146 %
2,3	1.069 %
1,6	.991 %

the correlation between rating and the percentage of rating count to installs



**THANK YOU  
FOR  
ATTENTION**

