# 微信用户画像

# 1 数据

#### 1.1 数据源

- 1. 微信用户月度宽表: cdg\_weixin\_secret::monthly\_wt\_weixin\_user\_attr\_info
- 2. 公众号协议
- 3. 阅读及文章行为协议
- 4. 标签协议

# 1.2 数据类型:

数据类型包括四种:连续型、离散型、空间数据型、时序数据型(暂未包括)

#### 目前画像工具支持:

**连续型**:年龄、注册年限、发消息天数、朋友圈发表天数、朋友圈点赞天数、朋友圈评论天数、好友数、支付等级、群数、文章阅读篇数、文章分享转发、关注订阅号数、点击订阅号次数、点击订阅号个数、点击订阅号入口次数

**离散型**:性别、学历、设备、阅读兴趣\_文章数均值、阅读兴趣>4篇用户数、转发兴趣文章数均值、转发兴趣>4篇用户数

空间型数据:居住城市

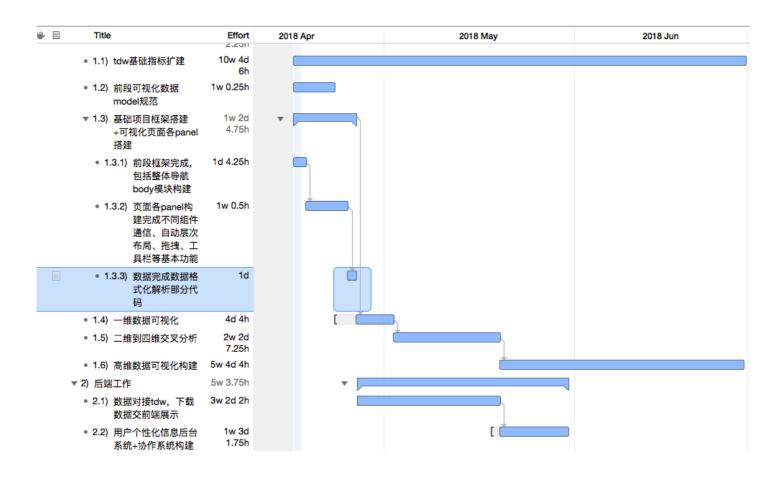
## 未来画像工具希望支持:

- 1、区分内链外链相关数据
- 2、用户活跃度指标实际定位到条数(现以天数为指标)

#### 其他

对于不同数据类型需要设计不同的可视化方式,进行交叉时,不同的数据类型交叉时可视化方案也应不同

# 2 基本进度规划



# 3 数据可视化方案

## 3.1 分类:

#### 3.1.1 不带参数的分布数据(用户画像):

用户画像主要展示为维度+相应人口分布,方法主要使用1~N维相应的可视化方案即可

#### 3.1.2 带参数的分布数据(试验系统):

带参数的分布相当于需要展示的值维度+1,由于试验系统中参数为连续型,因此,相当于天然的连型维度+1,再按相应的方案聚合即可。

# 3.2 方案:

#### 一维

#### 图形:

连续型: 折线图
 离散型: 柱状图

3. 地理数据型: 地图热力图或城市散点

#### 数据接口:

1. 连续型: 折线图

2. 离散型: 柱状图 (堆叠、柱状)

```
{
   title:"用户学历分布",
   group: ["user1", 'user2', 'user3'],
   label: {
              attr0:"未知",
              attr1:"本科",
              attr2:"研究生"
          },
   data:[
       [[attr0,3],[attr1,2],[attr2,1]], // [x,y] x: 相同值只能出现一次,多次出现
只取最后一次
       [[attr0,1],[attr2,2],[attr1,5]], // [x,y] x: 顺序可不必排序,组件可视化前
自动排序
       [[attr0,4],[attr1,2],[attr2,6]],
   ]
}
```

3. 地理数据型:城市散点图(多用户包分离的地图展示)

```
{name: "招远", value: 12},
            {name: "舟山", value: 12},
            {name: "齐齐哈尔", value: 14},
            {name: "盐城", value: 15},
            {name: "赤峰", value: 16},
            {name: "青岛", value: 18},
            {name: "乳山", value: 18},
            {name: "金昌", value: 19},
            {name: "泉州", value: 21},
            {name: "莱西", value: 21},
            {name: "日照", value: 21},
            ],
            {name: "海门", value: 9},
            {name: "鄂尔多斯", value: 12},
            {name: "招远", value: 12},
            {name: "舟山", value: 12},
            {name: "齐齐哈尔", value: 14},
            {name: "盐城", value: 15},
            {name: "赤峰", value: 16},
            {name: "青岛", value: 18},
            {name: "乳山", value: 18},
            {name: "金昌", value: 19},
            {name: "泉州", value: 21},
            {name: "莱西", value: 21},
           {name: "日照", value: 21},
            ],
            [
            {name: "海门", value: 9},
            {name: "鄂尔多斯", value: 12},
            {name: "招远", value: 12},
            {name: "舟山", value: 12},
            {name: "齐齐哈尔", value: 14},
            {name: "盐城", value: 15},
            {name: "赤峰", value: 16},
            {name: "青岛", value: 18},
            {name: "乳山", value: 18},
            {name: "金昌", value: 19},
            {name: "泉州", value: 21},
            {name: "莱西", value: 21},
            {name: "日照", value: 21},
            ]
    ]
}
```

#### 图形:

1. 连续型X连续型: 热力图

2. 离散型X离散型: 折线图

3. 连续型\*地理数据型:省份城市双编码,颜色+形状双编码4. 离散型\*地理数据型:省份城市双编码,颜色+形状双编码

#### 数据接口:

1. 连续型X连续型:点-热力图(多用户包分离的地图展示)

2. 离散型X离散型:块-热力图(多用户包分离的地图展示)

3. 连续型\*地理数据型: 省份城市双编码, 颜色+形状双编码

```
{
   title:"地理数据分布",
   group: ["user1", 'user2', 'user3'],
   label: {
               value1:"取值",
               value2:"人数",
           },
   data:[
           [
               {province: "海门",city: "海门", value1: 9 ,value2: 9},
               {province: "内蒙",city: "鄂尔多斯", value1: 12 ,value2: 12},
               {province: "浙江",city: "舟山", value1: 12 ,value2: 12},
               {province: "山东",city: "青岛", value1: 18, value2: 18},
           ],
           [
               {province: "海门",city: "海门", value1: 9 ,value2: 9},
               {province: "内蒙",city: "鄂尔多斯", value1: 12 ,value2: 12},
               {province: "浙江",city: "舟山", value1: 12 ,value2: 12},
               {province: "山东",city: "青岛", value1: 18, value2: 18},
           ],
           [
               {province: "海门",city: "海门", value1: 9 ,value2: 9},
               {province: "内蒙",city: "鄂尔多斯", value1: 12 ,value2: 12},
               {province: "浙江",city: "舟山", value1: 12 ,value2: 12},
               {province: "山东",city: "青岛", value1: 18, value2: 18},
           ],
   ]
}
```

4. 离散型\*地理数据型: 省份城市双编码, 颜色+形状双编码

```
{
   title:"地理数据分布",, 吗
                                  group: ["user1", 'user2', 'user3'],
   label: {
               value1:"取值",
               value2:"人数",
           },
   data:[
               {province: "海门",city: "海门", value1: 9 ,value2: 9},
               {province: "内蒙",city: "鄂尔多斯", value1: 12 ,value2: 12},
               {province: "浙江",city: "舟山", value1: 12 ,value2: 12},
               {province: "山东",city: "青岛", value1: 18 ,value2: 18},
           ],
               {province: "海门",city: "海门", value1: 9 ,value2: 9},
               {province: "内蒙",city: "鄂尔多斯", value1: 12 ,value2: 12},
               {province: "浙江",city: "舟山", value1: 12 ,value2: 12},
               {province: "山东",city: "青岛", value1: 18, value2: 18},
           ],
               {province: "海门",city: "海门", value1: 9 ,value2: 9},
               {province: "内蒙",city: "鄂尔多斯", value1: 12 ,value2: 12},
               {province: "浙江",city: "舟山", value1: 12 ,value2: 12},
               {province: "山东",city: "青岛", value1: 18, value2: 18},
           ],
   ]
}
```

#### 三维

#### 多图矩阵:

- 1. 将连续性分段化为离散型,找到维度最少的属性,对图进行分块,从而相应的可视化矩阵
- 2. 可引入体数据可视化在三维空间进行刻画

#### 多视觉通道映射:

通道包括: 颜色、形状、大小、透明度、位置等, 可尝试在基础图上通过多通道进行映射。

#### 高维

#### 交叉矩阵:

将高维数据向两个属性维度进行投影,从而形成维度矩阵,利用缩略与交互的方式,通过缩略找特征,然后 利用交互查看详细信息。

# 降维与推荐:

- 1. 分析给定数据主元,只展示主元维度信息。
- 2. 利用机器学习方法找特征,从而进行低维度交叉推荐。
- 3. 用户群体划分类别,先找类别再反过来刻画属性
- 4. 其他

# 附录:

# 一、数据分布于来源

# 1 连续型变量:

#### 年龄

划分与取值:

数据来源: fage

#### 注册年限

范围划分与取值:

```
注册年限 微信MAU
1 164498
2 163336
3 181144
4 135813
5 152940
6 138327
>=7 70601
```

数据来源: fopen\_time (注册时间戳)

## 发消息天数

发消息天数 微信MAU
0 19679
1~5 141439
6~10 80754
11~15 78106
16~20 89174
21~25 122780
26~31 474727

数据来源: fsendmsg\_days

#### 朋友圈发表天数

范围划分与取值:

朋友圈发表天数 微信MAU
0天 354231
1~5天 344373
6~10天 125199
11~15天 71419
16~20天 46295
21~25天 33012
26~31天 32130

数据来源:fsnsupload\_days

#### 朋友圈点赞天数

范围划分与取值:

朋友圈点赞天数 微信MAU
0 357355
1~5 255335
6~10 111394
11~15 82823
16~20 69190
21~25 61419
26~31 69143

数据来源: fsnspraise\_days

## 朋友圈评论天数

# 朋友圈评论天数 微信MAU 0 401434 1~5 242708 6~10 115677 11~15 82186 16~20 64584 21~25 53013 26~31 47057

数据来源: fsnscomm\_days

#### 好友数

范围划分与取值:

```
好友数 微信MAU
0 13744
1~20 169824
21~50 157760
51~100 179302
101~200 206128
201~400 162451
401~800 77812
≥801 39638
```

数据来源: ftotal friend count

#### 支付等级

范围划分与取值:

```
支付等级
         微信MAU
未知 427978
lv 1
       12
lv 2
      159
lv 3
      1485
lv 4
      14563
lv 5
      61787
lv 6
      209430
lv 7
       291245
```

#### 群数

```
群数 微信MAU
0 117073
1~5 358038
6~10 210730
11~20 197654
21~40 96477
41~100 23986
101~200 2183
≥201 518
```

#### 文章阅读篇数

范围划分与取值:

```
文章阅读篇数 微信MAU
0 155437
1~5 165892
6~20 220014
21~50 186109
51~100 127342
101~200 90347
201~400 44883
≥401 16635
```

# 文章分享转发

范围划分与取值:

```
文章分享转发 微信MAU
0 338616
1~5 284600
6~20 218499
21~50 100589
51~100 38388
≥101 25967
```

#### 数据来源:

## 关注订阅号数

```
      关注订阅号数 微信MAU

      0个 129325

      1个 78739

      2~5个 205756

      6~10个 161538

      11~20个 186495

      21~50个 185936

      51~100个 49621

      101~200个 8455

      201~500个 776

      500+个 18
```

#### 点击订阅号次数

#### 范围划分与取值:

```
点击订阅号次数 微信MAU
0次 194483
1次 44734
2~5次 152036
6~10次 102797
11~20次 109051
21~50次 154786
51~100次 105838
101~200次 77022
201~500次 50394
500+次 15518
```

#### 点击订阅号个数

#### 范围划分与取值:

```
点击订阅号个数 微信MAU
0个 194483
1个 170574
2~5个 280263
6~10个 161668
11~20个 118728
21~50个 67068
51~100个 11663
101~200个 1934
201~500个 270
500+个 8
```

#### 点击订阅号入口次数

#### 范围划分与取值:

```
点击订阅号入口次数 微信MAU
0次 224159
1次 40292
2~5次 121230
6~10次 99783
11~20次 121676
21~50次 180908
51~100次 120499
101~200次 72417
201~500次 24656
500+次 1039
```

## 2 离散型变量:

#### 性别

范围划分与取值:

```
性别 微信MAU
男 518782
女 477771
```

数据来源: fsex

#### 学历

范围划分与取值:

```
学历 微信MAU
其他 366541
博士 5261
硕士 15592
本科 247889
高中 133511
初中 125315
小学 85803
大专 26747
```

数据来源: fgrade (1:博士,2:硕士,3:本科,4:高中,5:初中,6:小学,7:大专)

#### 设备

设备 微信MAU 其他 19958 iOS 203738 Android 782963

数据来源: flastuseddevice

#### 阅读兴趣 文章数均值

#### 范围划分与取值:

阅读兴趣 文章数均值 比例 微信MAU 图片 0.6505329181776313 视频 0.2593850220688015 投票 0.008332865278979118 长文章 0.5970063707690478 动漫 0.14729431935504864 音乐 0.015883279838975688 影视 0.06828430583887672 收藏 0.006444463238825345 星座占ト 0.02151333778518784 彩票 0.04480332015632675 植物 0.006700564014006909 茶艺 0.006252141638786873 文学 0.025469299146705194 五金 0.023178169214599843 爱心公益 0.029716917941453317 节假日 0.03623284727784075 鸡汤故事 0.014829280230191762 旅游 0.04134867982213501 佛教 0.017038472025497558 书画 0.006371660285190392 摄影 0.005702600955610715 人文地理 0.009319613600319583 基督教 0.004407057448883127 历史 0.0344916665127116 中国传统文化 0.03493112081177614 球类 0.008769627298452468 健身健美 0.017525180473248907 政法类新闻 0.10445866566352947 财经类新闻 0.06576869549120223 交通类新闻 0.04570298715847621 治安 0.04511412062582318 法律 0.02969651231889662 国际新闻 0.04476661231665924 污染 0.005270462230056034

救援 0.0185331328173007

能源产业 0.02256941694786289

农产品 0.015729355716141265

社保养老公积金 0.01659201779949591

食品安全 0.015197918292285199 医药管理 0.011883740682971917

金融 0.03634635007191636

健康险 0.010142355675942745

住宅 0.022615464212595023

企业管理 0.06995212662700723

互联网 0.019739942502191656

手机 0.011939127372768669

教育 0.06059584232191104

研究生院 0.015464101190349645

本科教育 0.027980323335351925

专科学校 0.05506647563240951

出国 0.010784993530637816

外语 0.0028270049754255218

家庭教育 0.012422865029246976

装修 0.01047127797220472

家居卫生 0.02173508878627487

各地菜系 0.057697129872132505

烟酒 0.005298573342477181

情感类资讯 0.22989518398528122

疾病 0.03542562758661922

养生 0.06157139447364124

孕产育儿 0.011331414882977283

汽车 0.03302088345576552

求职招聘 0.023290298017711673

职场生涯 0.03489994607085467

场所 0.03490521922536432

购物资讯 0.08734035362256895

端游 0.012883801862704696

其他游戏 0.012947042581929549

美容护理 0.015989299952532326

化妆 0.015238358188489466

整形手术 0.00875644441217835

防晒 0.06175218569005122

服装搭配 0.019210231850763045

#### 阅读兴趣\_>4篇用户数

#### 范围划分与取值:

阅读兴趣 >4篇用户数 占比 微信MAU

图片 0.6376091605995675

视频 0.49827598024753167

投票 0.020580951444332193

长文章 0.6062291202880022

```
动漫 0.3055125916521881
音乐 0.056522615900717124
影视 0.17755367011073264
收藏 0.017570994745986477
星座占ト
         0.06574023577000752
彩票 0.16899466452890205
植物 0.021327977001149346
茶艺 0.020488566634779007
文学 0.08588906471804256
五金 0.07999034429732411
爱心公益
         0.11385980754158062
节假日 0.1438163270779877
鸡汤故事
        0.05700639441956015
旅游 0.14714416699200028
佛教 0.051164296946632375
书画 0.017564041050643762
摄影 0.016940195239897524
人文地理
        0.03302607933769032
基督教 0.006860317148110731
历史 0.08970366330604504
中国传统文化 0.11128396010962997
球类 0.024355814630376324
健身健美 0.06300147319002761
政法类新闻 0.2616615954359917
财经类新闻
          0.17554802569688444
交通类新闻
          0.16235587224670917
治安 0.16030751227575574
法律 0.11136144414344877
国际新闻
         0.11844825308272215
污染 0.016254759556115824
救援 0.07384327761436593
能源产业
         0.07212670824976482
农产品 0.05677493570315271
社保养老公积金 0.0651014891835269
食品安全
        0.057808056154070045
医药管理
         0.04393146040516203
金融 0.11163661180201041
健康险 0.03529000386426784
住宅 0.07938338603241019
企业管理 0.19972801117359504
互联网 0.059188861372123036
手机 0.035133049026532324
教育 0.19353723554848265
研究生院
        0.05366663388496005
本科教育
         0.10093288789947738
专科学校
        0.178464604200628
```

出国 0.038194661747423905 外语 0.008702053028880684 家庭教育 0.04367516706253061

装修 0.034116816121447284

家居卫生 0.08133340088351666

各地菜系 0.17959010946109855

烟酒 0.016009393449022954

情感类资讯 0.44298218165237685

疾病 0.12217742055651418

养生 0.186698772871449

孕产育儿 0.04114600872788104

汽车 0.1069458476008261

求职招聘 0.08505561466196597

职场生涯 0.12901886338869467

场所 0.12997847334598905

购物资讯 0.28820583732922467

端游 0.03776949294646946

其他游戏 0.04516822479111596

美容护理 0.05544876666279246

化妆 0.04377450556742651

整形手术 0.02775517826791396

防晒 0.17260959272206378

服装搭配 0.052484505676698864

#### 转发兴趣\_文章数均值

#### 范围划分与取值:

转发兴趣 文章数均值 比例 微信MAU

图片 0.13547367417072625

视频 0.06141271309704748

投票 0.001148756883097259

长文章 0.123613501145207

动漫 0.016088959368661275

音乐 0.002908228316954288

影视 0.006867893025422818

收藏 0.0013166062650016223

星座占卜 0.003805546022210773

彩票 0.012742085560050293

植物 0.00143047057036372

茶艺 0.0017100409295292475

文学 0.00640814914777256

五金 0.0038340120985512977

爱心公益 0.005245527611022386

节假日 0.013381199208147433

鸡汤故事 0.003328454582743583

旅游 0.008814838689108383

佛教 0.006395891990195346

书画 0.0016332829919145626

摄影 7.425292289671865E-4

人文地理 0.0016550511679396696

基督教 0.0017491566673712858

历史 0.00563909623355632

中国传统文化 0.010562856713426048

球类 6.139965219143242E-4

健身健美 0.0049786832624561525

政法类新闻 0.019164501235494694

财经类新闻 0.011884955256686314

交通类新闻 0.007085239890658117

治安 0.006170507644581546

法律 0.005777809749088617

国际新闻 0.0070299822130559225

污染 0.0014187492448117394

救援 0.002224640610762776

能源产业 0.004136623234801861

农产品 0.002854377198418331

社保养老公积金 0.003987996826802747

食品安全 0.003714722493933712

医药管理 0.0029311351360330157

金融 0.007511427287728134

健康险 0.003111576570530364

住宅 0.004289536299002843

企业管理 0.01396539007365815

互联网 0.003690744010804517

手机 0.0013075640995758085

教育 0.01425078760609809

研究生院 0.003906952232986195

本科教育 0.005924493765996261

专科学校 0.012889774262005249

出国 0.0024381026938151322

外语 7.512364993772292E-4

家庭教育 0.0035877303039533954

装修 0.0018963095373010087

家居卫生 0.004468839090446571

各地菜系 0.01328247215749818

烟酒 9.144643300639543E-4

情感类资讯 0.057724447309359675

疾病 0.01090665993661643

养生 0.01630657414990919

孕产育儿 0.0030920187016093446

汽车 0.0035709855531648516

求职招聘 0.004406280701500571

职场生涯 0.010601905472264932

场所 0.0059084188052392584

购物资讯 0.015203496946963078

端游 6.395825011192191E-4

其他游戏 0.001736698572784609

美容护理 0.0031685087232114126

化妆 0.001150632295185576

整形手术 0.0010125885696848208

防晒 0.007699504328585058

服装搭配 0.001726316827295712

#### 转发兴趣 >4篇用户数

#### 范围划分与取值:

转发兴趣 >4篇用户数 占比 微信MAU 图片 0.11501809450866679 视频 0.05688619482863611 投票 7.142438502015082E-4 长文章 0.1056455065717388 动漫 0.015827603985063464 音乐 0.002260944371430643 影视 0.005735805272689163 收藏 9.715305778818846E-4 星座占ト 0.002782471522134109 彩票 0.012125257907593336 植物 9.407356413641561E-4 茶艺 0.001233784230807056 文学 0.005467591309470238 五金 0.002907638038302941 爱心公益 0.0037480417897222396 节假日 0.012916985791613645 鸡汤故事 0.0022559774461858484 旅游 0.007820920490454067 佛教 0.004806990251912515 书画 0.0012437180812966456 摄影 4.1026802522005964E-4 人文地理 0.0011970289839955735 基督教 0.0012576254719820713 历史 0.004699704666624944 中国传统文化 0.008767616442111976 球类 3.4768476713564376E-4 健身健美 0.004485133496049804 政法类新闻 0.017375297891341557 财经类新闻 0.010646107569693412 交通类新闻 0.00570997726141623 治安 0.0047006980516739036 法律 0.004768248235003114 国际新闻 0.005880839489837174 污染 8.90073003867248E-4 救援 0.0013659044423186004 能源产业 0.0034182379534678574 农产品 0.0021983611133462276

社保养老公积金 0.0032453889549489945

食品安全 0.0030288310142759364 医药管理 0.0024119388988724086

金融 0.006525546386611554 健康险 0.0026453843853777694

住宅 0.00344505934978975

企业管理 0.012936853492592824

互联网 0.0030198905488353057

手机 8.245095906359551E-4

教育 0.013754409387886067

研究生院 0.00332783991401259 本科教育 0.0049788458653824185

专科学校 0.012854402533529229

出国 0.0016331250204885666

外语 5.572890124659889E-4

家庭教育 0.002843068010120607

装修 0.001234777615856015

家居卫生 0.003671551140952398 各地菜系 0.012705394776185382

烟酒 5.255006908993016E-4

情感类资讯 0.055605721500527984

疾病 0.010359019290544266

养生 0.015363693167199618

孕产育儿 0.0024606147662713986

汽车 0.0021367712403107704

求职招聘 0.003540424314489812 职场生涯 0.009914976173659601

场所 0.004582485230847784

购物资讯 0.014030570431496664

端游 3.4271784189084884E-4

其他游戏 0.0010281535256725465 美容护理 0.00260564898341941

化妆 8.503376019088887E-4

整形手术 6.953695342712875E-4

防晒 0.0064808440594084

服装搭配 0.0012923939486956357

## 3 空间型数据

#### 居住城市

居住城市 微信MAU

特级(北上广深) 83850

一线 231008

二线 198378

三线及以下 438803

国外 33277

数据来源: 注册城市-fregcity 最常登录城市-fcommonused\_city

# 二、协议整理

#### 11954 朋友圈分享协议

prePublishId curPublishId:

msg表示会话消息的id, sns表示朋友圈feed的id, fav表示收藏的id, gs表示微信全局搜索, bs*表示公众号搜索, bshw*表示公众号热词搜索, app\_ 从第三方app分享, custom*menu 自定义菜单,brand*profile 公众号 profile, jd*store 京东购物入口,pay*wallet钱包入口,game*center*gift 游戏中心礼包,game*center*entrance游戏中心h5,pay*blance*list零钱明细,wx*sport 微信运动, game*center 游戏native, music*player 音乐播放器, shake*music 摇音乐, scan*product 扫商品, shake*tv 摇TV, emoje*stroe 表情商店, card*package 卡包, wxapp*appid* 应用号, wechat\_out WeChat out