

01 背景摘要

目录 CONTENTS 02 数据采集

03数据存储

04 数据分析



背景摘要

突破关键技术

带动新兴学科

发展高新技术产业



战略科学家

科技创新创业领军人才

科技创新创业青年人才



数据采集

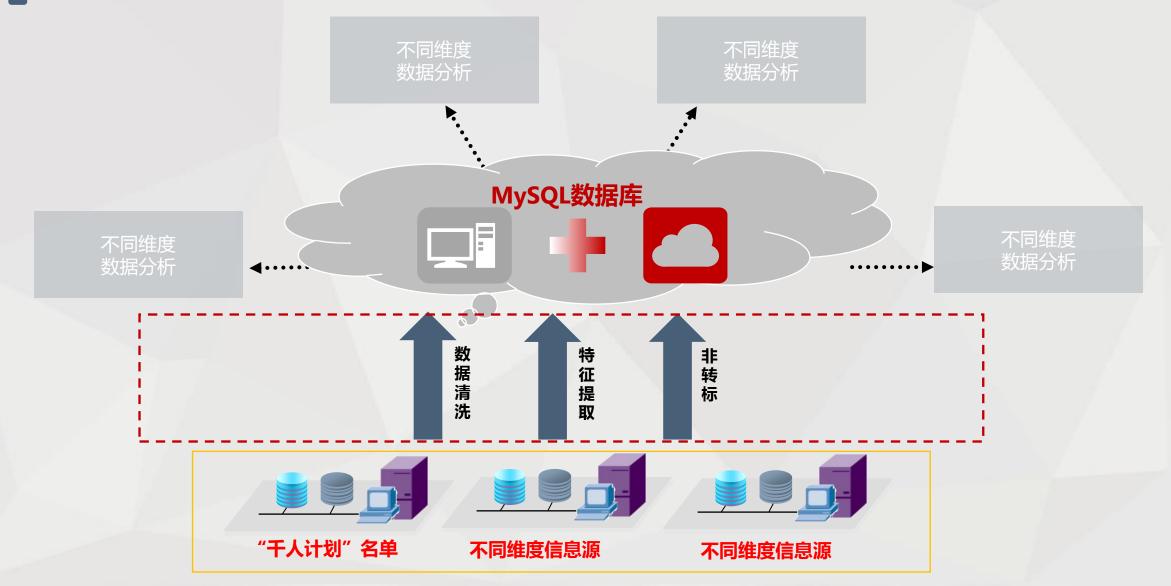


数据采集:

基于中组部发布的2016年度"干人计划"名单定向搜集"性别"、"年龄"、"毕业院校"等14个画像标签数据

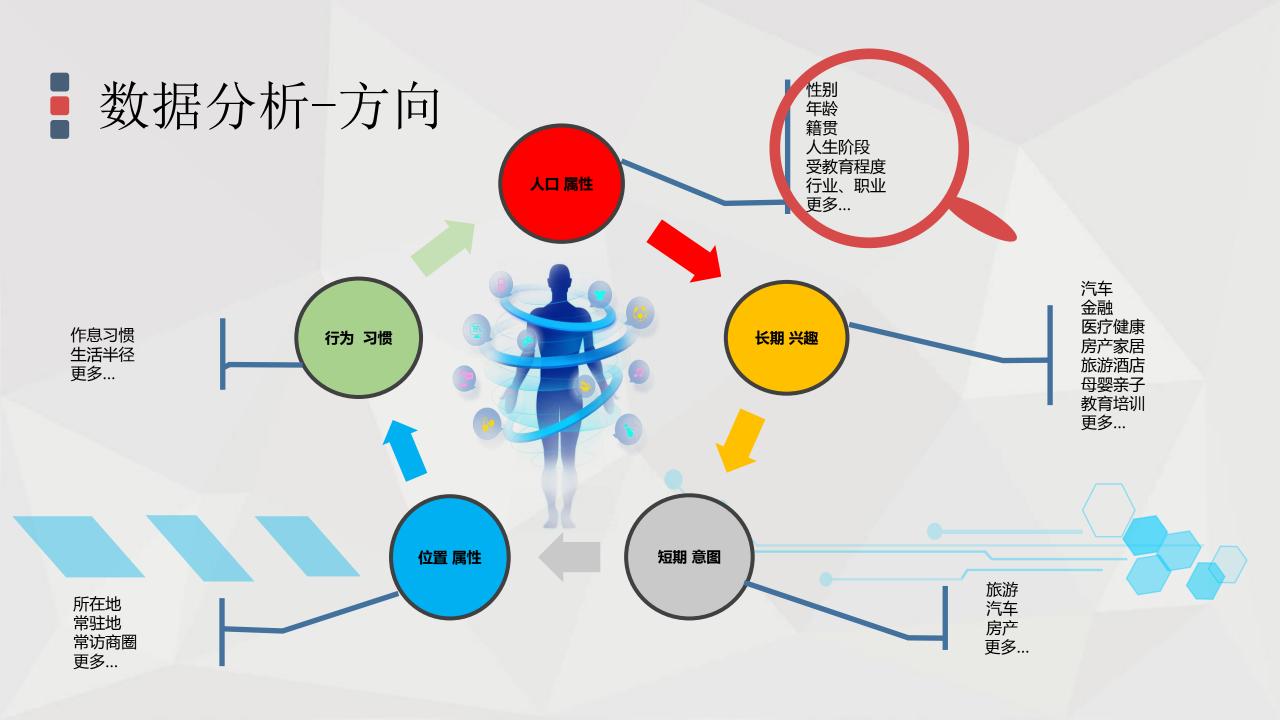












数据分析-结论











年龄

以1982年为中轴的类 正态分布

毕业院校

美国、中国为首的 各国名牌大学

专业学科

化学、物理生物、材料等 基础专业学科

性别

以男士为主

所属 / 获奖单位

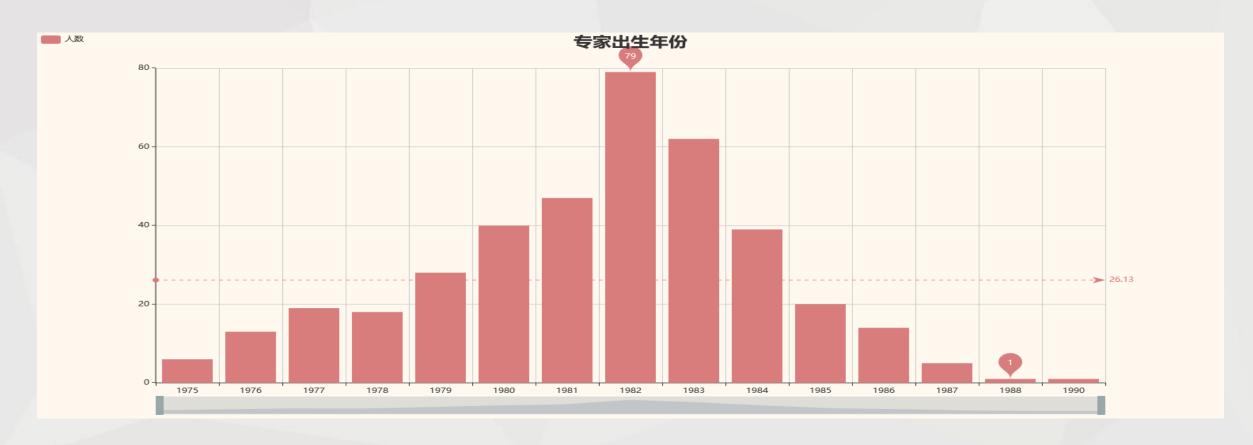
以名牌大学为主

年龄维度统计分析

```
mysql> select birth_year,count(*) total from(
    -> select substr(birthday,1,4) birth_year from thousand_talents_program
    -> )t
    -> group by birth_year
    -> order by total desc;
  birth_year | total
  1982
                  79
  1983
                  62
                  47
  1981
  1980
                  40
  1984
                  39
  1979
                  28
  1985
                  20
  1977
                  19
  1978
                  18
  1986
                  14
 1976
                  13
 1975
                   6
 1987
                   5
  1988
  1990
15 rows in set (0.00 sec)
```

年龄维度统计分析

- > 年龄近似成正态分布
- ▶ 1982年为峰值,人数向两侧递减



毕业院校维度统计分析

<pre>mysql> select school,count(name) total -> from thousand_talents_program -> where length(school)>1 -> group by school -> order by total desc;</pre>	
school	total
SCR001 ** *********************************	16 12 11 10 8 7 7 5 4 4 4 4 4 3 3 3 3 3
佐治亚理工学院 南达科他州立大学 生津大学	2 2 2

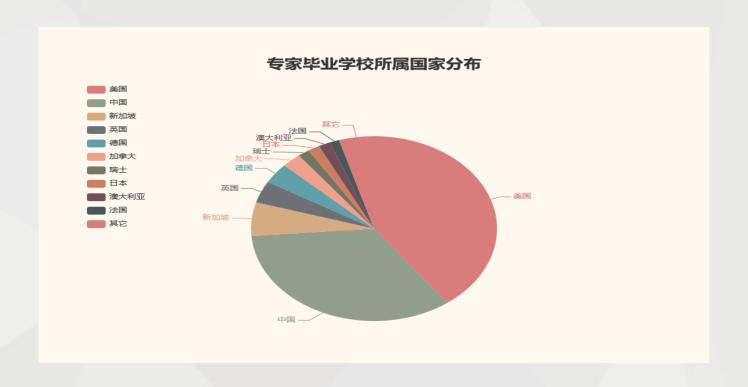
毕业院校维度统计分析





毕业院校属国维度统计分析

- ▶ 专家毕业院校以美国、中国为主
- > 大部分专家毕业于各国名牌大学



专业维度统计分析

```
mysql> select major_abbreviation,count(name) total
    -> from thousand_talents_program
    -> where length(major_abbreviation)>1
    -> group by major_abbreviation
-> order by total desc;
  major_abbreviation | total |
  化学
                         59
  物理
                         55
                         49
  生物学
  材料学
                         36
                         27
  医学
                         20
  数学
  环境科学
                         17
                          7776
  机械
  电子
  计算机
  自旨领原
  土木工程
                          6
  天文
                          4
                          4
  海洋科学
  光学
                          4
                          4
  高分子
                          3
                          3
  精密仪器
  植物学
  工学
                          33222222
  生态学
  通信
  农学
  自动化
  空间科学
  微电子
                          ī
  系统工程
                          ī
  软件
  经济学
                          ī
1
  药学
  汽车工程
  冶金
                          1
  纳米科学
37 rows in set (0.00 sec)
```

专业维度统计分析

> 专业以基础科学为主,包括化学、物理生物、材料等

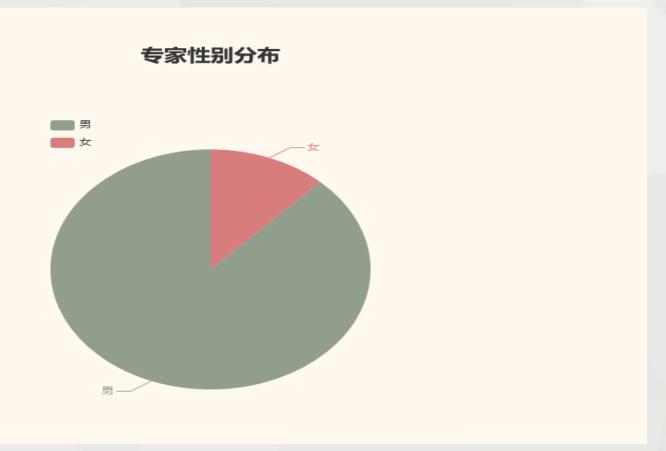


性别维度统计分析

```
mysql> select sex,count(name) from thousand_talents_program group by sex;
+----+
| sex | count(name) |
+----+
| 女 | 47 |
| 男 | 345 |
+----+
2 rows in set (0.00 sec)
```

性别维度统计分析

> 专家性别以男士为主



所属单位维度统计分析

mysql> select organization,count(name) total	
-> from thousand talents program	
<pre>-> where length(organization)>1</pre>	
-> group by organization	
-> order by total desc;	
+	
organization	total
+	
┃ 北京大学	28
中国科学技术大学	27
│ 上海交通大学	22
┃ 复旦大学	j 19 j
清华大学	18
中国科学院上海生命科学研究院	13
浙江大学	12
厦门大学	10
│ 南京大学	10
┃ 上海科技大学	j 9 j
同济大学	j 9 j
· 北京航空航天大学	j 8 j
, 京南大学	j 7 j
山东大学	j 6 j
华中科技大学	j 6 j
哈尔滨工业大学	6
, 南开大学化学学院	4
北京师范大学	4
华东师范大学	4
浙江大学生命科学研究院	3
中国科学院大气物理研究所	3
中国科学院国家纳米科学中心	3
上海大学	3
│ 东北大学流程工业综合自动化国家重点实验室	2
浙江大学电气工程学院	2
中国科学院苏州纳米技术与纳米仿生研究所	2
北京交通大学	2
长春应用化学研究所	2
浙江大学数学科学学院	2
南开大学物理科学学院	2
┃ 苏州大学医学部	2
南京工业大学化工学院	2
浙江大学材料科学与工程学院	2
北京工业大学	2

所属单位维度统计分析

> 所属单位以名牌大学为主



获得称号单位维度统计分析

▶ 获得称号单位以名牌大学为主





感恩谆谆教诲学生受益良多