



# 全球运维大会

2016

重新定义运维

上海站

会议时间： 9月23日-9月24日

会议地点： 上海·雅悦新天地大酒店

主办单位：  开放运维联盟  
OOPSA Open OPS Alliance  高效运维社区  
Great OPS Community

指导单位：  数据中心联盟  
Data Center Alliance



# 大数据时代运维的机遇与挑战

尹会生 西山居



# 个人介绍

尹会生



西山居运维经理  
重庆工商大学研究生导师  
《白话大数据与机器学习》作者之一

# 目录



**1** 用数据说话

**2** 数据可视化

**3** 用数据决策



# 1 用数据说话

## 1. 运维接触到的数据是什么样的？

- 各种各样的日志
- 价值密度相对较低

## 2. 它们有什么用吗？

- 用户行为的有力证据
- 系统可用性可靠性最直接的证明
- 监控
- 扩容



# 1 用数据说话

## 3. 如何处理这些日志呢？

- Hadoop ( HDFS )
- Flume

## 4. 我们做个高大上的入口

- Nginx
- Python



# 1 用数据说话

## 5. 给业务

- 业务
- 计划

The screenshot displays the Hive Query Editor interface. The top navigation bar includes 'Query Editors', 'Notebooks', 'Data Browsers', 'Workflows', 'Search', and 'Security'. The main workspace shows a SQL query for computing the total amount per order for all customers, filtered by zip code '94123'. The query is as follows:

```
18 -- Compute total amount per order for all customers
19 SELECT
20 c.id AS customer_id,
21 c.name AS customer_name,
22 ords.order_id AS order_id,
23 SUM(order_items.price * order_items.qty) AS total_amount
24 FROM
25 customers c
26 LATERAL VIEW EXPLODE(c.orders) o AS ords
27 LATERAL VIEW EXPLODE(ords.items) i AS order_items
28 GROUP BY c.id, c.name, ords.order_id;
```

The execution results show a single row for zip code '94123'. Below the query, the 'Query History' section lists several recent queries, including the one shown in the main workspace.



# 目录

1 用数据说话

➔ 2 数据可视化

3 用数据决策

# 2 数据

## 1. 我们

- Data
- BI

DATANOSE

- 实时概况
- 报表
- 玩家分析
- 付费分析
- 流失分析
- 在线分析
- 渠道分析
- 系统分析
- 自定义事件分析
- 运营支持
- 数据挖掘
- 市场分析
- 技术支持
- 管理中心

DATANOSE

下午 02:15:07

设备激活: 实时统计当日激活的设备数量, 如果已安装的游戏  
 新增玩家: 实时统计当日新增玩家账号的数量。  
 老玩家: 实时统计当日有登陆过游戏的老玩家的账号数量。  
 总活跃: 实时统计当日所有进行过登陆行为的玩家账号数量

设备激活(含重复安装)	1000 +1000	总活跃	1000 +1000
新增玩家	1000 +1000	老玩家	1000 +1000

实时在线 设备激活 新增玩家 活跃玩家

2016-09-01 添加 清空

人数

12/14 12/15

DATANOSE

demo@datanose.com

2016-09-01 刷新 2016-09-01 今日 昨日 前日 前7日 前30日 前90日

设备激活: 统计所激活的设备, 每日新增玩家安装游戏设备, 并进行游戏的可连接设备数量的统计, 每台设备只计算一次。  
 新增用户: 统计所激活的设备, 每日玩家首次登陆, 进行了自动或者手动注册和创建账号或用户设备的游戏用户数量。  
 老玩家: 统计所激活的设备, 每日玩家重新登陆, 进行了自动或者手动注册和创建账号或用户设备的游戏设备数量, 每台设备只计算一次。  
 活跃玩家: 统计所激活的设备, 每日玩家重新登陆, 进行了自动或者手动注册和创建账号或用户设备的游戏设备数量, 即每日登录并注册的设备数量, 如果设备注册不在当日进行注册则不计入。

新增玩家 玩家活跃度

SUM(新增玩家设备): 91210 AVG(活跃玩家设备): 91210

日期	设备	新增用户	新增设备
2015-01-01	11	0	21
2012-01-01	11	10	21
2018-01-01	1	0	2

1 2 3 4 5 6 7 8 9 10

首次登陆时长 小等级分析 地区 国家 性别 年龄 用户类型

日期	设备	新增用户	新增设备
2015-01-01	11	0	21
2012-01-01	11	10	21
2018-01-01	1	0	2

1 2 3 4 5 6 7 8 9 10

sol = sol =



项目1 算法 图表

算法模块: k-均值 密度聚类 层次聚类

创建图表任务

任务名称: 项目1\_chart\_2\_22

x轴: AGE

y轴: USER

\*y轴只能选择数值型字段, 其他类型会导致图表无法生成

我的数据源

项目1 算法 图表

数据库

Customer

PRODUCT

SINGULAR

我的首页

我的项目

我的报表

我的收藏夹

我的任务

我的数据源

运算任务列表

已完成运算任务 未完成运算任务

操作	任务名称	项目名称	开始时间	结束时间	进度	状态	发起人	任务耗时 (天)
查看结果 删除任务	k-均值	项目1	2015-10-01	2015-10-01	100%	完成	Jonny	0.5

批量删除

1

SINGULAR

运算任务列表

已完成运算任务 未完成运算任务

操作	任务名称	项目	进度	状态	发起人	任务耗时 (天)
查看结果 删除任务	k-均值	项目1	100%	完成	admin	0.5
查看结果 删除任务	密度聚类	项目1	100%	完成	admin	0.3

批量删除

1

查看任务结果

文件名	缩略图	大小 (kb)	操作
k-均值图片		50kb	预览 下载
k-均值文本		15kb	查看 下载

# 目录

1 用数据说话

2 数据可视化

➔ 3 用数据决策

# 3 用数据决策



Prisma



点击或拖拽图片到这里

选择你要生成的图片风格



生成图片

solz





# Thanks

高效运维社区  
开放运维联盟

**荣誉出品**





## 想第一时间看到高效运维公众号的好文章么？

请打开高效运维公众号，点击右上角小人，并如右侧所示设置即可：



# GOPS2016 全球运维大会更多精彩

## GOPS2016 全球运维大会·北京站

2016年12月16日-17日  
北京国际会议中心

