

BI产品中的可视化原理 及应用实践

刘阳

字节跳动 DataWind可视化负责人





写作平台

- ☑ InfoQ 写作平台是 InfoQ 开放给**开发者的高端技术社区**，创作者可以在这里**自由创作和发布内容**。
- ☑ 写作平台将为创作者**提供签约、培训、资金扶持**等一系列权益，助力作者成长为高精尖技术人才；同时也为企业**提供品牌、活动打造、内容传播**等服务，与伙伴一同成长。

扫码申请创作者
企业/个人均可申请



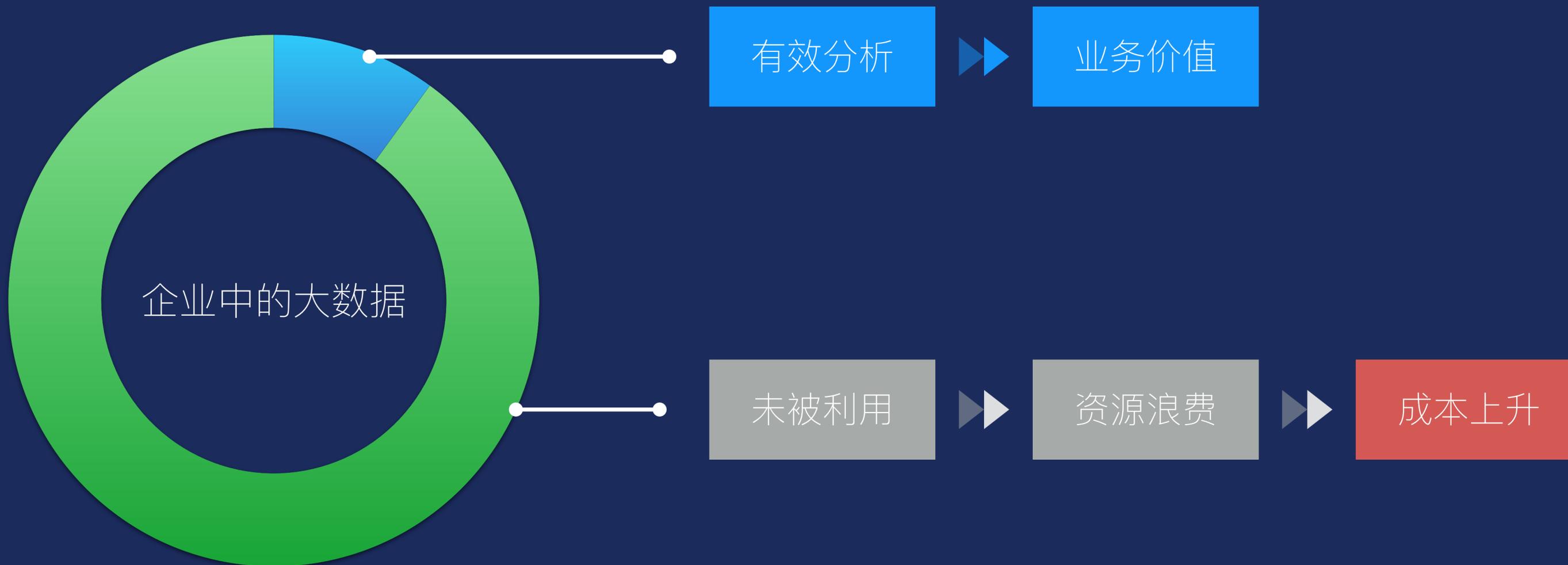
扫码进入写作平台
企业/个人均可申请



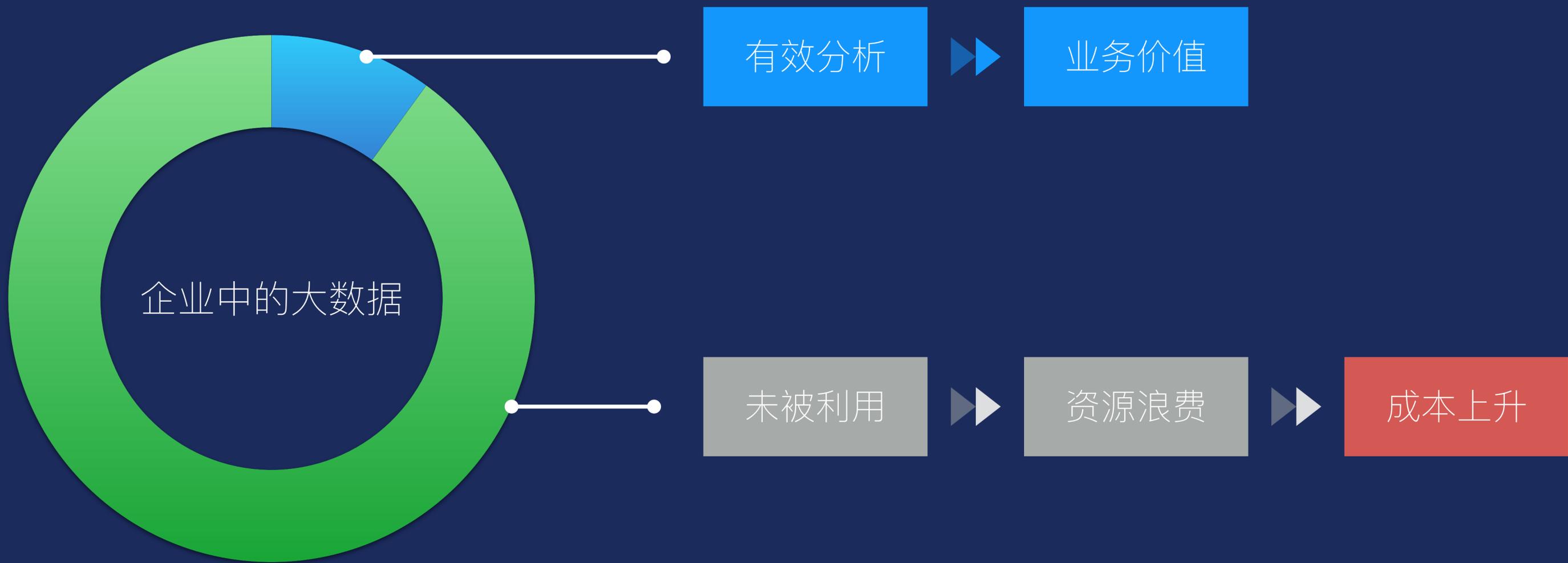
大纲

- BI的概念与背景
- BI系统开发中常遇到的问题
- 基于可视化图形语法的实践
- 数据可视化的技术演进
- 总结

BI (Business Intelligence)



BI的概念与背景



有超过80%的企业其数据以非结构化为主，超过90%的企业内部存在数据孤岛，约80%的企业不认可自身数据挖掘能力，同时仅有不到40%的企业采购第三方数据，多数企业没有对外寻求优质、合规的第三方数据供应商的意识。

《2019中国企业数字化转型及数据应用调研报告》

传统型BI vs 敏捷型BI

- 部署成本
- 报表灵活性
- 大数据分析能力
- 数据响应速度

BI的概念与背景

敏捷型BI系统中的可视化流程



连接数据源



数据编排



发布仪表盘

BI的概念与背景

BI平台架构图



大纲

- BI的概念与背景
- **BI系统开发中常遇到的问题**
- 基于可视化图形语法的实践
- 数据可视化的技术演进
- 总结

BI系统开发中常遇到的问题

- 产研团队间的协作沟通
- 可视化呈现效果
- 可视化的回归测试
- 可视化需求无法穷举
- 千奇百怪的数据和用法
- 大数据量的性能

BI系统开发中常遇到的问题

- 产研团队间的协作沟通
- 可视化呈现效果
- 可视化的回归测试
- 可视化需求无法穷举
- 千奇百怪的数据和用法
- 大数据量的性能

专业性

BI系统开发中常遇到的问题

- 产研团队间的协作沟通
- 可视化呈现效果
- 可视化的回归测试
- 可视化需求无法穷举
- 千奇百怪的数据和用法
- 大数据量的性能

数据与可视化的分离

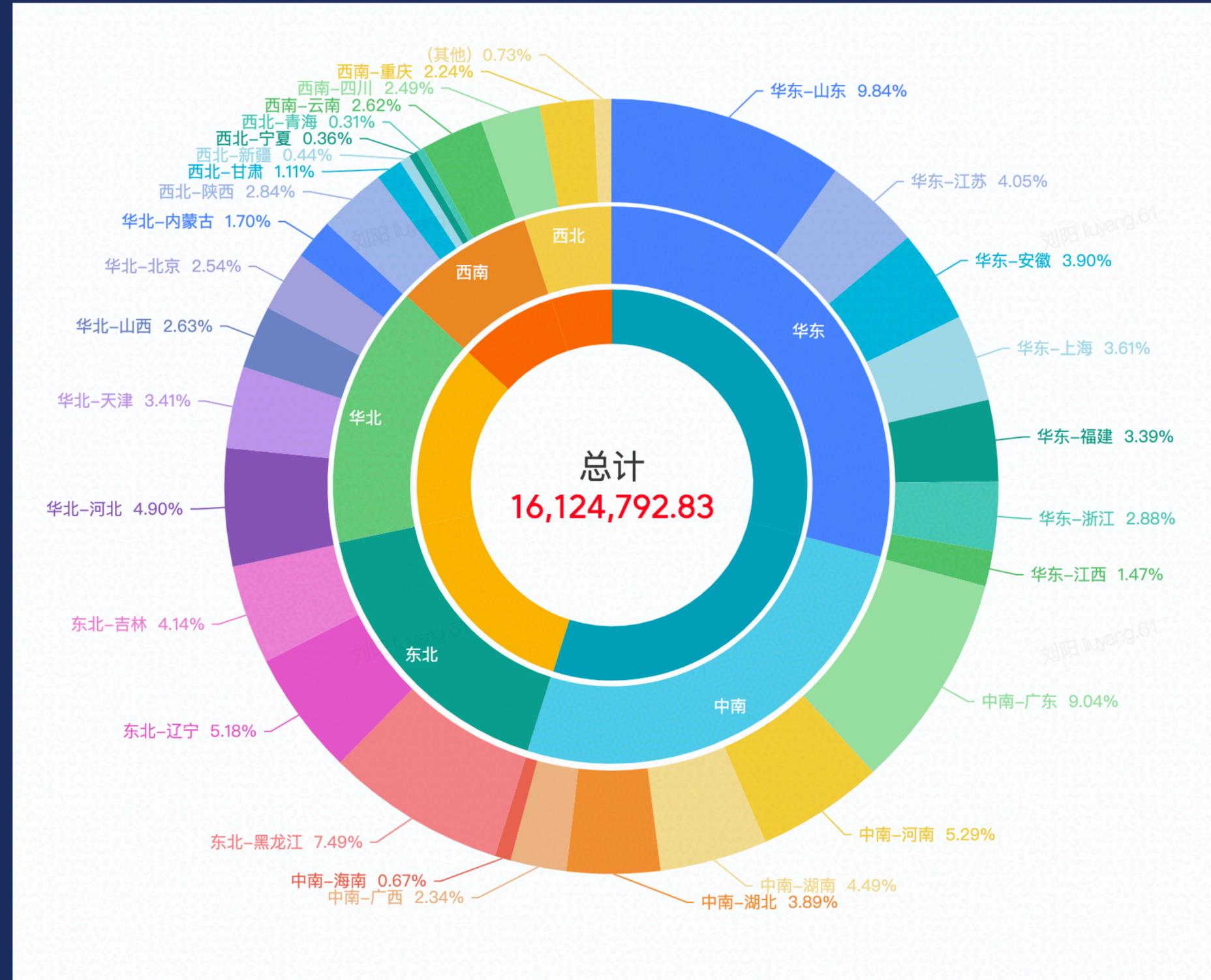
开脑洞的用法

BI系统开发中常遇到的问题

- 产研团队间的协作沟通
- 可视化需求无法穷举

数据与可视化的分离

开脑洞的用法



问题

- 产研团队间的协作沟通
- 可视化呈现效果
- 可视化的回归测试
- 可视化需求无法穷举
- 千奇百怪的数据和用法
- 大数据量的性能

数据与可视化的分离

开脑洞的用法

BI系统开发中常遇到的问题

- 产研团队间的协作沟通
- 可视化呈现效果
- 可视化的回归测试
- 可视化需求无法穷举
- 千奇百怪的数据和用法
- 大数据量的性能

图形测试

功能的排列组合

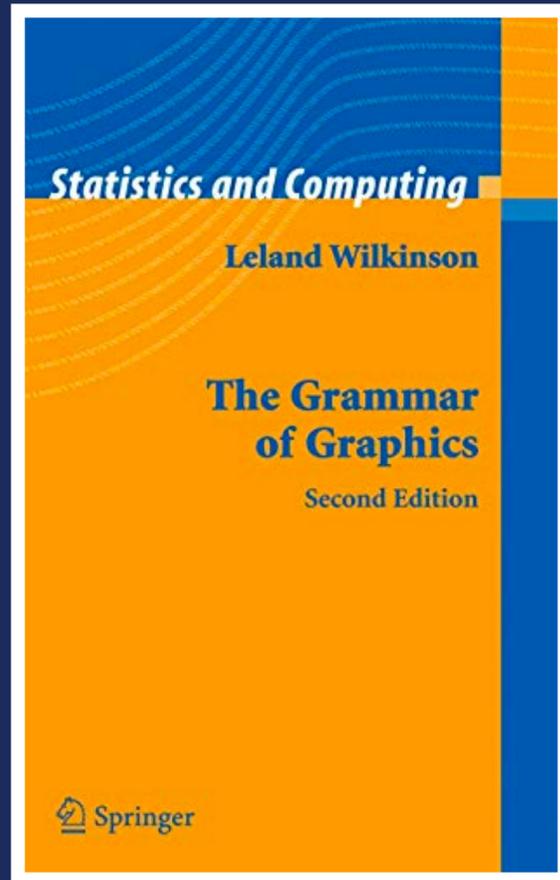
海量数据需要秒级响应

BI系统开发中常遇到的问题

大纲

- BI的概念与背景
- BI系统开发中常遇到的问题
- **基于可视化图形语法的实践**
- 数据可视化的技术演进
- 总结

可视化图形语法



《The Grammar of Graphics》

- 图表元素自由组合
- 图表元素高度自治
- 元素属性数据驱动

基于可视化图形语法的实践

BI中实践图形语法

一切行为都是对数据字段的处理

BI中的功能



抽象为字段行为



影响查询取数



图表展示

基于可视化图形语法的实践

BI中实践图形语法

举例：上卷下钻



基于可视化图形语法的实践

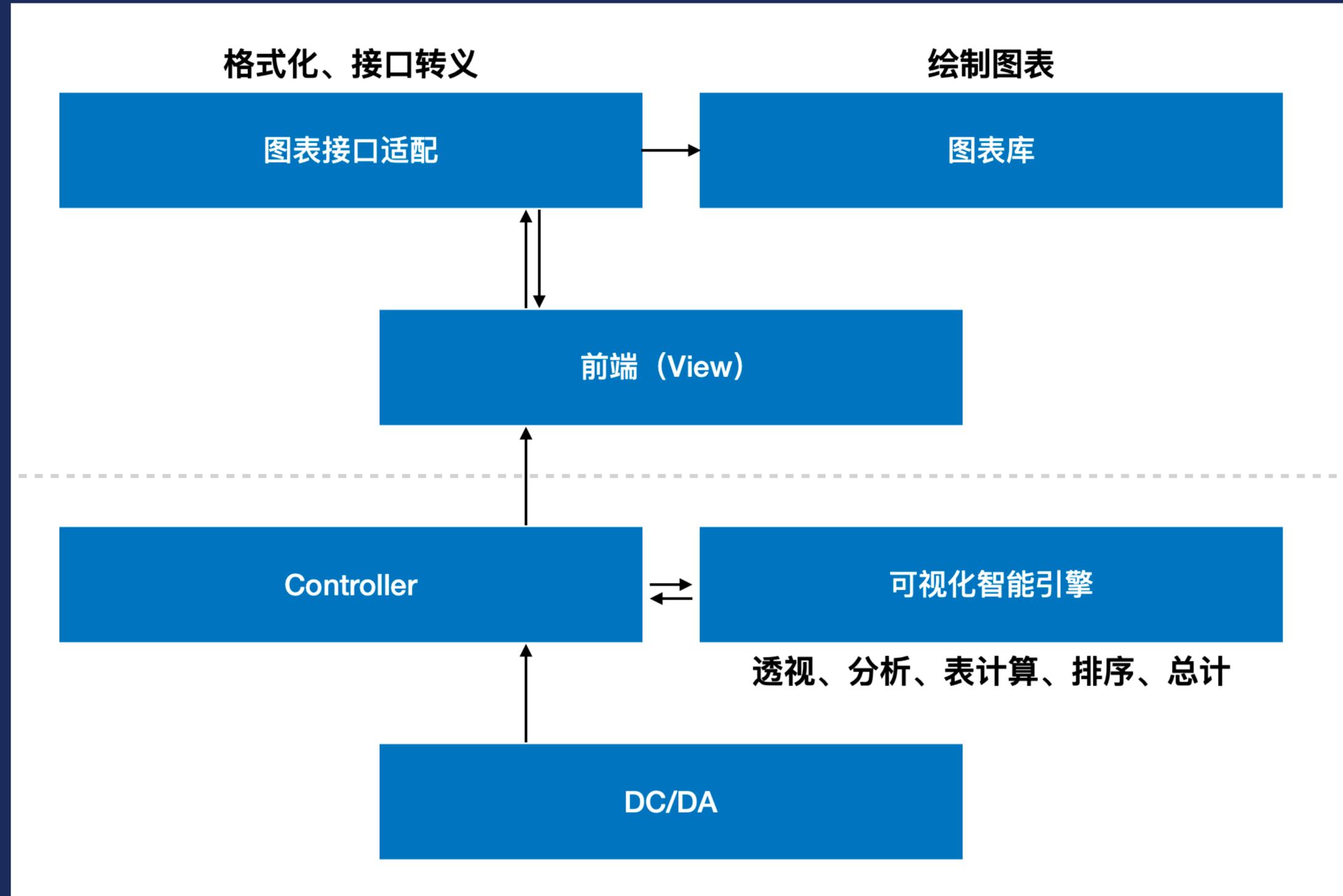
字节跳动的数据系统实践

- 可视化架构设计
- 基于图形语法的图表库
- 图表回归测试

基于可视化图形语法的实践

可视化架构设计

- 可视化智能引擎
- columns to columns
- 用户行为解析





ChartSpace

- 引擎层基于图形语法
- 声明式语法
- 自研绘制底层
- 开箱即用

基于可视化图形语法的实践

版本: feat_pivot_richtext 9bc3 图片缩放自适应: OFF 渲染方式: canvas webgl 标准图表对比

测试用例: 静态渲染

测试图表 测试时间: 2021-06-11 18:11 替换

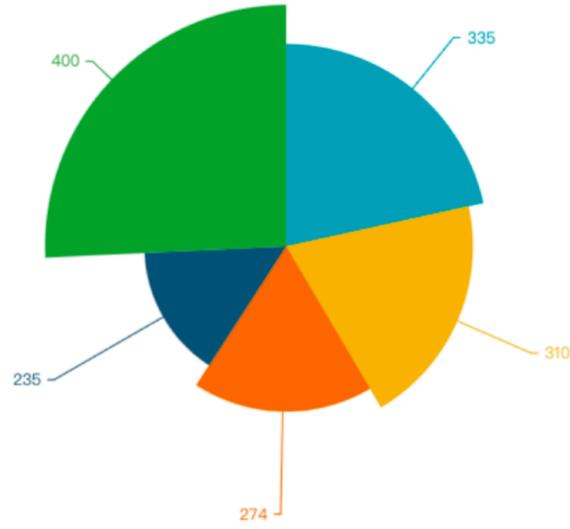
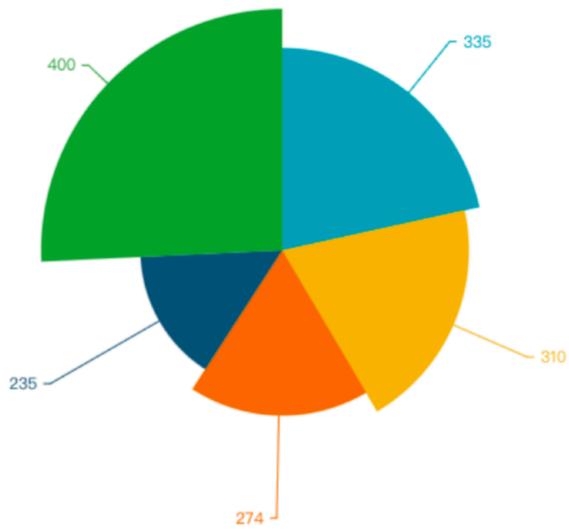
标准图表

差异对比 成功

直接访问 邮件营销 联盟广告 视频广告 搜索引擎

直接访问 邮件营销 联盟广告 视频广告 搜索引擎

直接访问 邮件营销 联盟广告 视频广告 搜索引擎



spec 更新时间: 2021-05-07 17:03 复制 编辑

```

1  {
2    "width": 500,
3    "data": [...],
29  ],
30  "type": "pie",
31  "radius": 0.8,
32  "nameField": "name",
33  "valueField": "value",
34  "groupBy": "name",
35  "legend": {...},
38  },
39  "labels": {...},
41  },
42  "roseType": "area"
43  }

```

未通过的测试/总计(0/0)

canvas webgl 对比

测试报告版本 test-sc test-s

对比新增错误 feat_pivot_richtext 9bc342

新增错误测试报告: 0 有差异的测试报告: 0

- bugSpec
- area
- axes
- bar
- box-plot
- combination
- datazoom
- funnel
- histogram
- hover
- indicator
- labels
- legend
- line
- map
- markline
- pie

版本: branch-2.13.1-alpha.2 b 图片缩放自适应: OFF 渲染方式: canvas webgl 标准图表对比

测试用例:

测试图表 测试用例已更新 测试时间: 2021-03-09 19: 27 替换

标准图表

差异对比 像素校验失败 错误率: 3.2733%



spec

更新时间: 2021-04-13 11: 47

复制

编辑

1

未通过的测试/总计(16/338)

canvas webgl 对比

测试报告版本 fix_map_data b031f0 1

对比新增错误 fix_map_data db5148 1

新增错误测试报告: 0 有差异的测试报告: 0

datazoom-line

canvas 静态渲染
错误率: 100.00% 标准图片获取失败

datazoom-line

canvas 交互 hover
错误率: 100.00% 标准图片获取失败

histogram

canvas 静态渲染
错误率: 100.00% 标准图片获取失败

histogram

canvas 交互 hover
错误率: 100.00% 标准图片获取失败

line-horizontal

canvas 静态渲染
错误率: 100.00% 标准图片获取失败

line-horizontal

canvas 交互 hover

大纲

- BI的概念与背景
- BI系统开发中常遇到的问题
- 基于可视化图形语法的实践
- **数据可视化的技术演进**
- 总结

数据可视化的技术演进

- Web Worker
- WebGL
- WebAssembly

数据可视化的技术演进

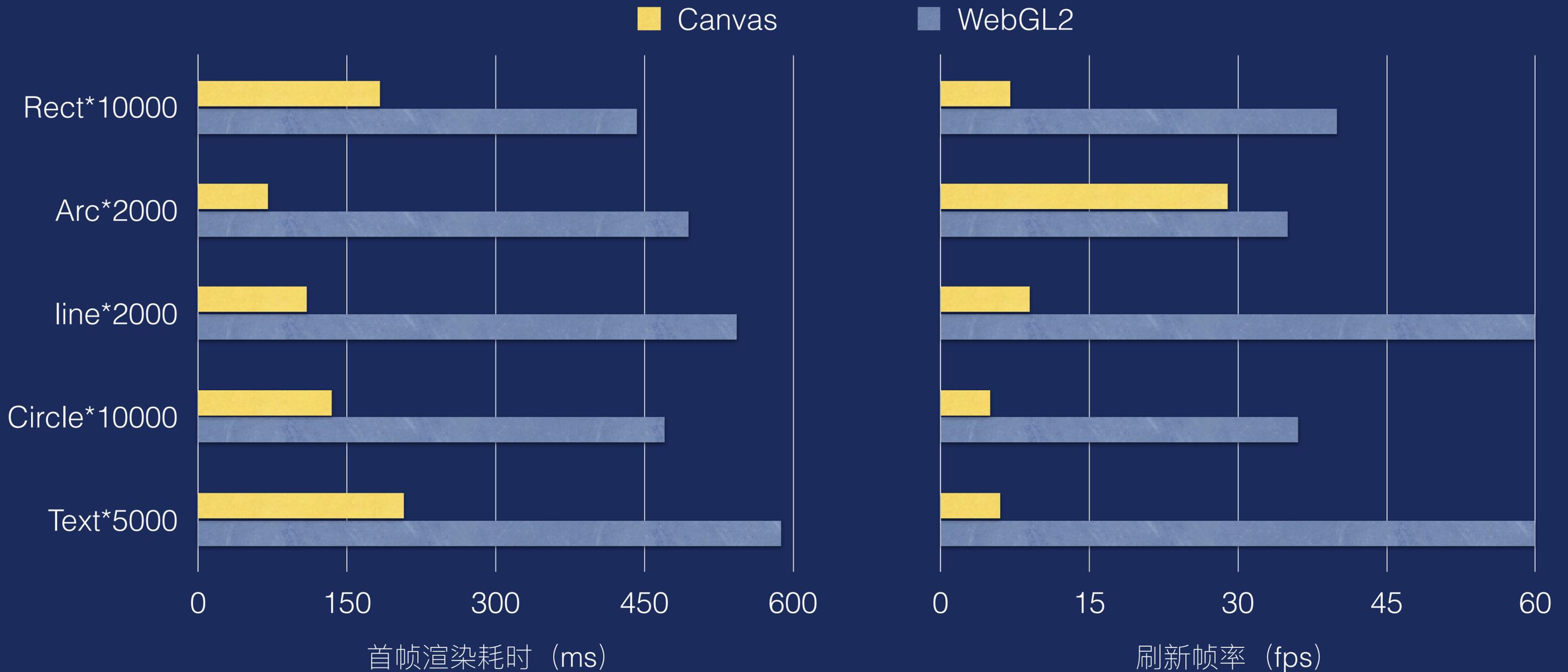
Web Worker

- 图表并行渲染
- OffscreenCanvas
- WorkerPool (线程池调度)

数据可视化的技术演进

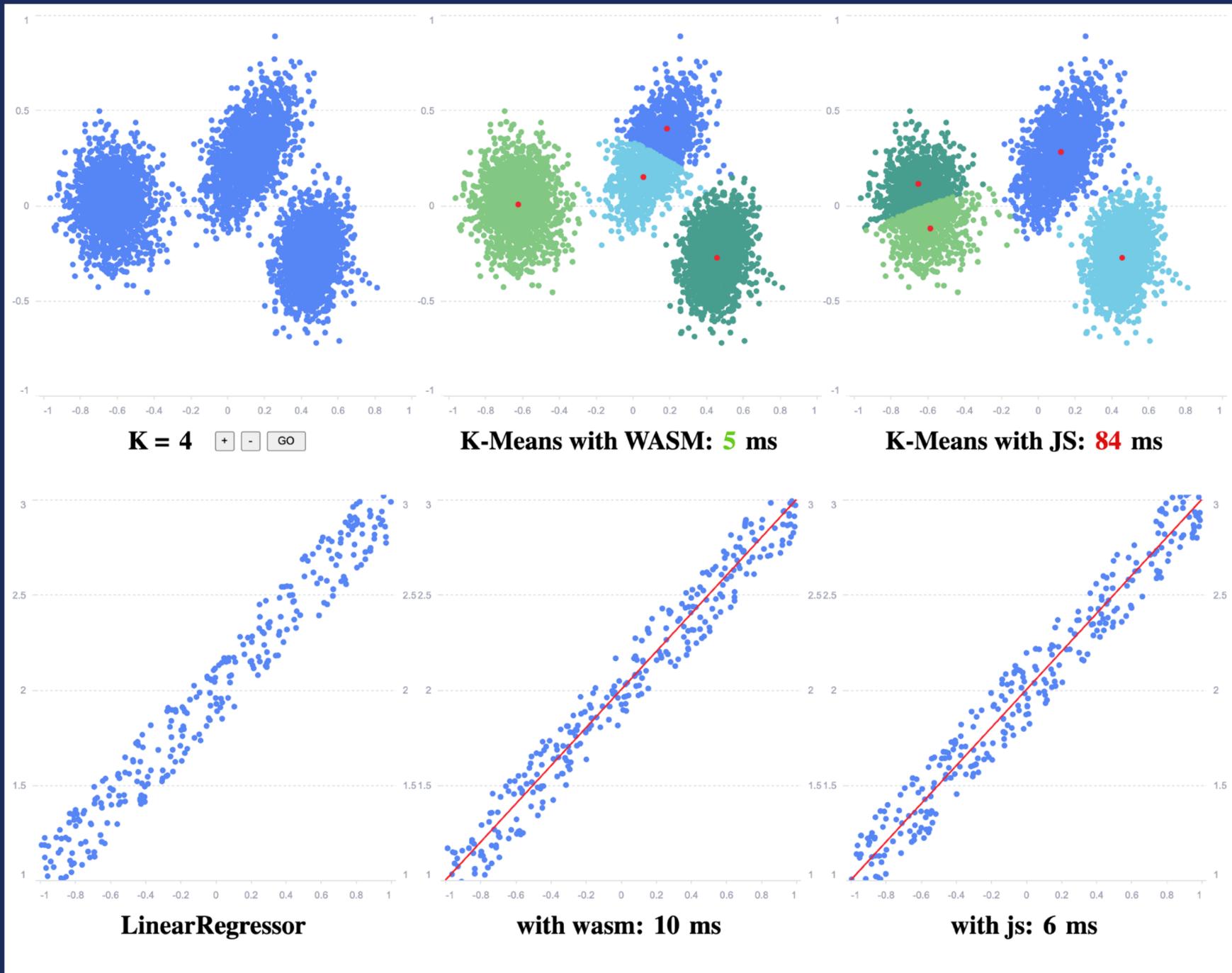


WebGL



WebAssembly

- wasm带来的性能提升相当可观
- wasm更适合复杂的计算逻辑
- 打包后的体积略大



数据可视化的技术演进

数据服务

- 数据解读服务
- 智能问答+数据洞察
- 流式数据与时空间数据

数据可视化的技术演进

大纲

- BI的概念与背景
- BI系统开发中经常遇到的问题
- 基于可视化图形语法的实践
- 数据可视化的发展趋势
- **总结**

总结

- 研发与学术结合
- 可视化 => 洞察 => 决策
- 商业价值

总结

GMTC
全球大前端技术大会

THANKS



WeChat: lyssfzx
字节跳动-刘阳



对标阿里 P7，带你打破 前端职业天花板

【前端实战训练营】3 个月带你进大厂 | 7月7日开营



扫码获取详细大纲
并咨询课程详情



—
8 大模块
教学



—
9 大实战
项目



—
15 周全程
直播授课



—
大厂助教
1v1 答疑



—
简历直推一线
互联网公司