

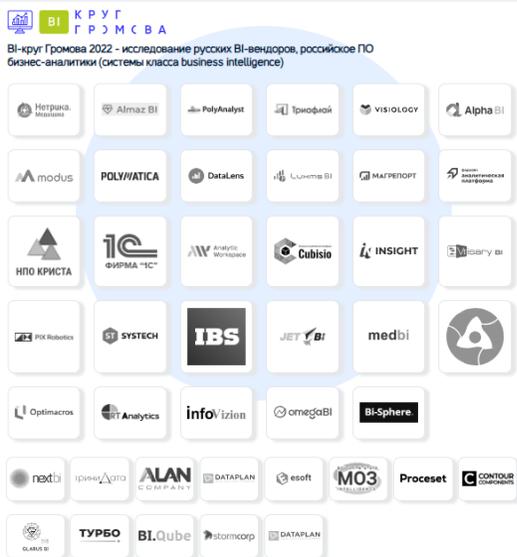


<https://datafinder.ru/>

Metabase



<https://biconsult.ru/>



<https://russianbi.ru/>
“Круги Громова” с 2019

К Р У Г И
Г Р О М О В А

Open-source BI

Плюсы:

- Экономически выгодная
- Более настраиваемая
- Поддержка сообщества
- Постоянные улучшения и инновации
- Нет вендор-лока
- Прозрачность кода

Минусы:

- Сложное внедрение
- Нет формальной поддержки
- Мало документации
- Обслуживание и upgrade
- Вопросы к совместимости и интеграции

Технические моменты

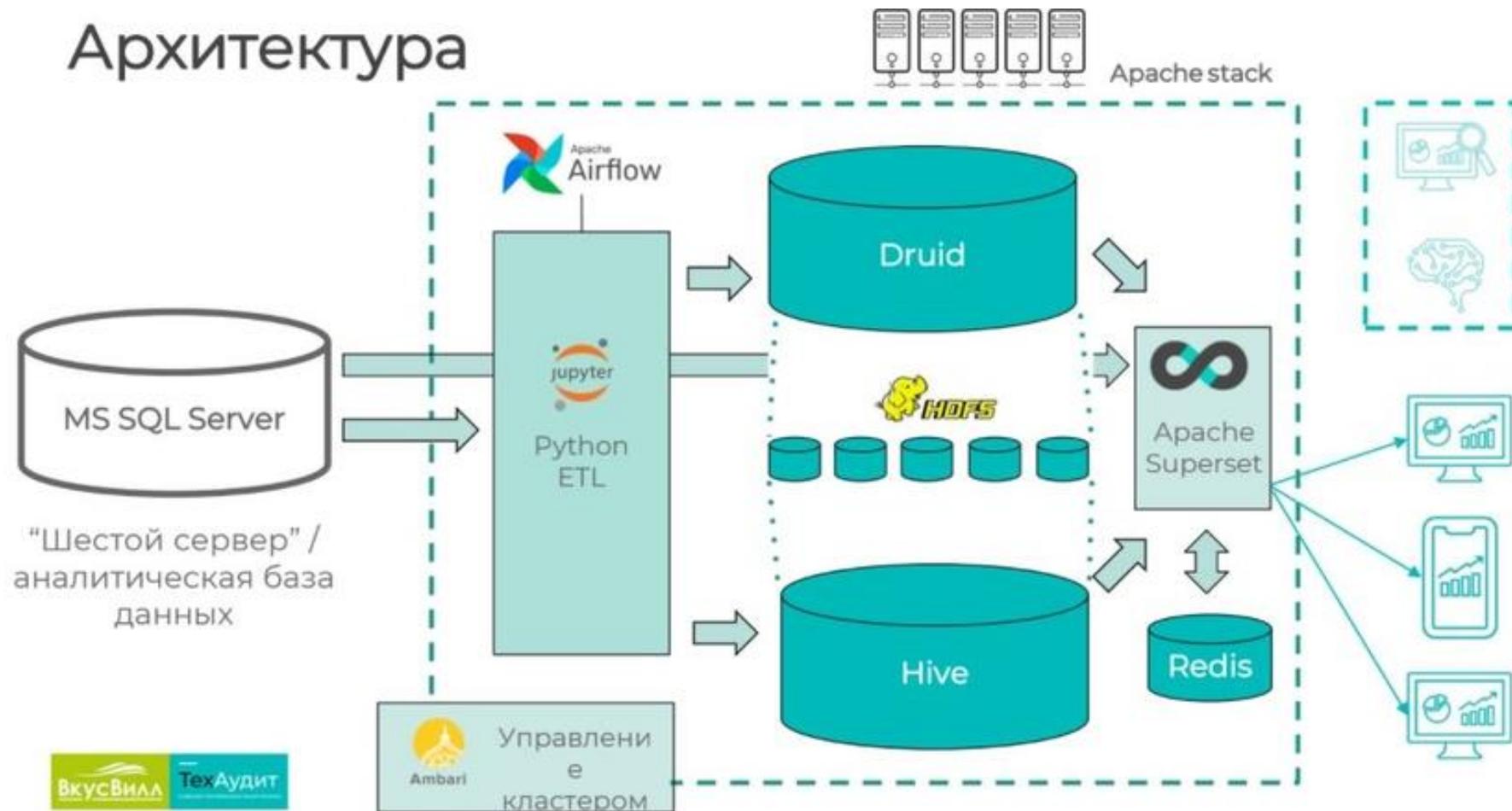
Плюсы

- Много визуализаций, настроек, фильтров
- Много типов источников
- Мультиязычность
- On-premise установка (на серверы компании)
- Интеграция с основными механизмами аутентификации (база данных, OpenID, LDAP, OAuth)
- Embed-доступ к дашборду по публичной ссылке
- Система уведомлений

Минусы

- Нужен опытный DevOps и программисты
- Работа с кодом
- Компоненты не оттестированы, могут не вставать
- Нефункциональный экспорт данных
- Могут случаться баги

Типовая архитектура решения Superset, на примере Вкусвилл (внедрение Техаудит)



Источник

[https://www.tadviser.ru/index.php/%D0%9F%D1%80%D0%BE%D0%B5%D0%BA%D1%82:%D0%98%D0%B7%D0%B1%D0%B5%D0%BD%D0%BA%D0%B0_\(%D0%92%D0%BA%D1%83%D1%81%D0%92%D0%B8%D0%BB%D0%BB\)_ \(Apache_Superset\)](https://www.tadviser.ru/index.php/%D0%9F%D1%80%D0%BE%D0%B5%D0%BA%D1%82:%D0%98%D0%B7%D0%B1%D0%B5%D0%BD%D0%BA%D0%B0_(%D0%92%D0%BA%D1%83%D1%81%D0%92%D0%B8%D0%BB%D0%BB)_ (Apache_Superset))

All charts

Search all charts

Recommended tags ^

Popular

ECharts

Advanced-Analytics

Category ^

Correlation

Distribution

Evolution

Flow

KPI

Map

Part of a Whole

Ranking

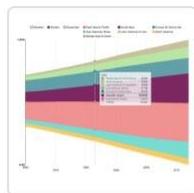
Table

Tags ^

2D

3D

Additive



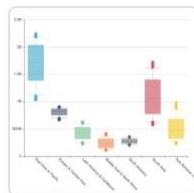
Area Chart



Big Number with Trendline



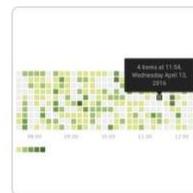
Big Number



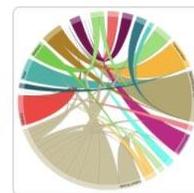
Box Plot



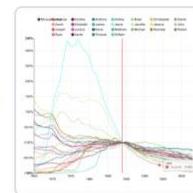
Bubble Chart



Calendar Heatmap



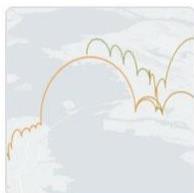
Chord Diagram



Time-series Percent Change



Country Map



deck.gl Arc



deck.gl Geojson



deck.gl Grid



deck.gl 3D Hexagon



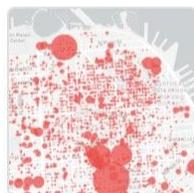
deck.gl Multiple Layers



deck.gl Path



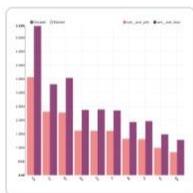
deck.gl Polygon



deck.gl Scatterplot



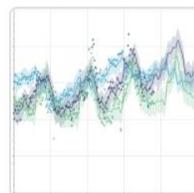
deck.gl Screen Grid



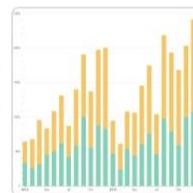
Bar Chart



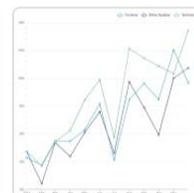
Time-series Area Chart



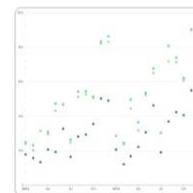
Time-series Chart



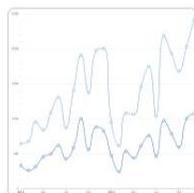
Time-series Bar Chart v2



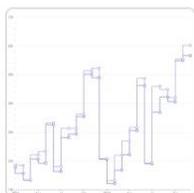
Time-series Line Chart



Time-series Scatter Plot



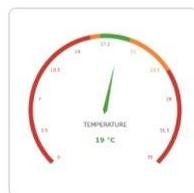
Time-series Smooth Line



Time-series Stepped Line



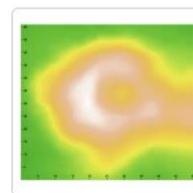
Funnel Chart



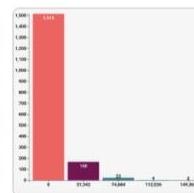
Gauge Chart



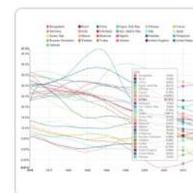
Graph Chart



Heatmap



Histogram



Line Chart

CLEAR ALL

APPLY

Channels

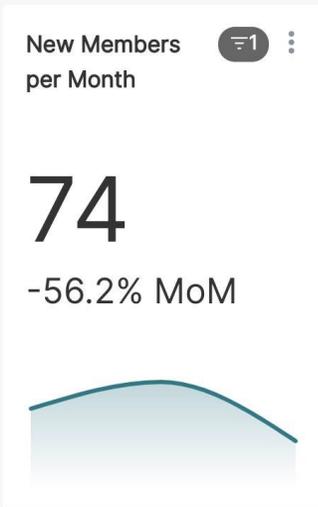
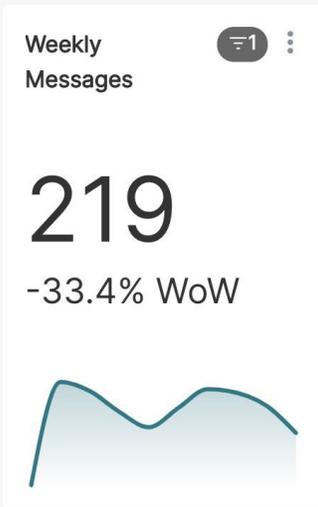
64 options 

Date

2020-09-13 ≤ col < 2020-11-13

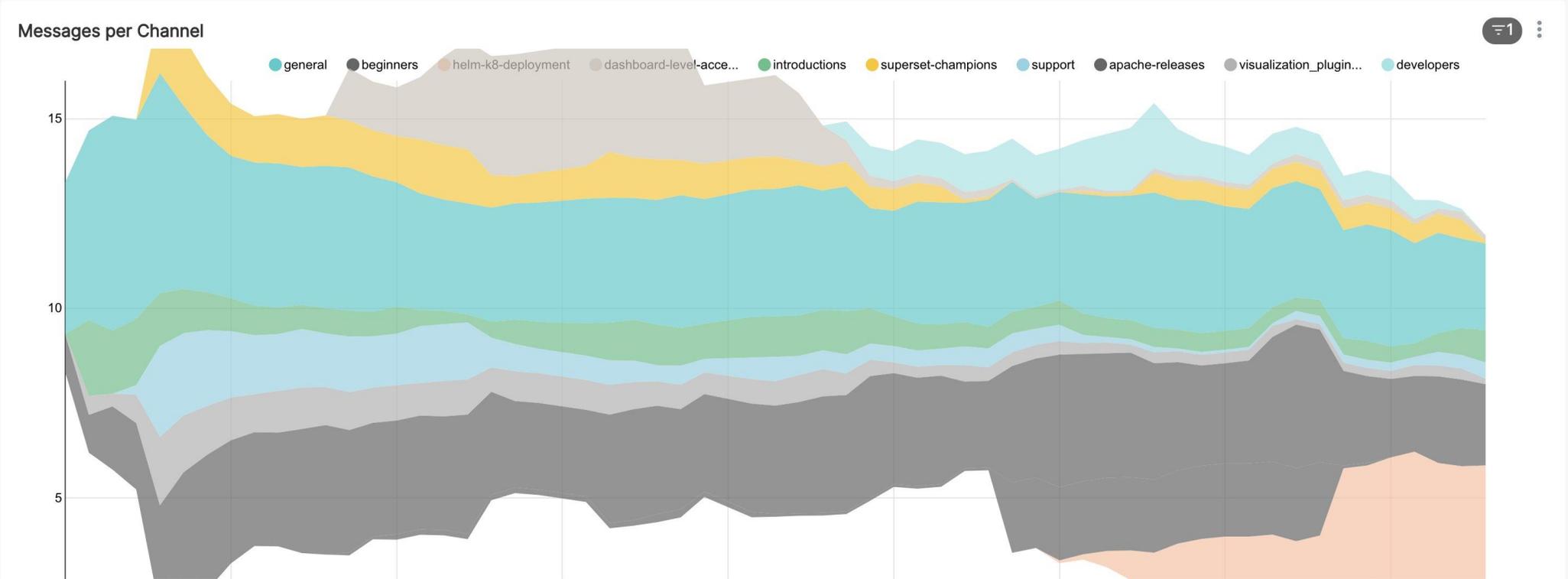


Number of Members
412
Slack Members



Top Timezones for Members

tz	COUNT(*)
America/Los_Angeles	60
Asia/Kolkata	57
America/New_York	48
Europe/Amsterdam	41
Asia/Chongqing	29
Europe/London	21
America/Chicago	18
Europe/Brussels	13
America/Sao_Paulo	12
America/Denver	11



Untitled Query 2

Database: postgresql examples

Schema: public

SEE TABLE SCHEMA 77 IN PUBLIC

Select table or type table name

covid_vaccines

- Developer_or_Researcher TEXT
- Country_Name TEXT
- IOC_Country_Code TEXT
- Treatment_vs_Vaccine TEXT
- Product_Category TEXT
- Stage_of_Development TEXT
- Date_Last_Updated TEXT
- Anticipated_Next_Steps TEXT
- Product_Description TEXT
- Clinical_Trials TEXT
- Funder TEXT
- Published_Results TEXT
- Clinical_Trials_for_Other_Diseases_or_Related_Use TEXT
- FDA_Approved_Indications DOUBLE PRECISION
- Sources TEXT

video_game_sales

- Rank BIGINT
- Name VARCHAR
- Platform VARCHAR
- Year BIGINT
- Genre VARCHAR
- Publisher VARCHAR
- NA_Sales DOUBLE PRECISION
- EU_Sales DOUBLE PRECISION
- JP_Sales DOUBLE PRECISION
- Other_Sales DOUBLE PRECISION
- Global_Sales DOUBLE PRECISION

```

1 SELECT "Rank",
2        "Name",
3        "Platform",
4        "Year",
5        "Genre",
6        "Publisher",
7        "NA_Sales",
8        "EU_Sales",
9        "JP_Sales",
10       "Other_Sales",
11       "Global_Sales"
12 FROM public.video
13 LIMIT 100
    
```

RUN

LIMIT: 1000

00:00:00.23

SAVE AS

COPY LINK

...

RESULTS QUERY HISTORY PREVIEW: 'COVID_VACCINES' PREVIEW: 'VIDEO_GAME_SALES'

EXPLORE

.CSV

CLIPBOARD

Filter Results

Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales
1	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	8.46
2	Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.77
3	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	3.31
4	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	2.96
5	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1
6	Tetris	GB	1989	Puzzle	Nintendo	23.2	2.26	4.22	0.58
7	New Super Mario Bros.	DS	2006	Platform	Nintendo	11.38	9.23	6.5	2.9
8	Wii Play	Wii	2006	Misc	Nintendo	14.03	9.2	2.93	2.85
9	New Super Mario Bros. Wii	Wii	2009	Platform	Nintendo	14.59	7.06	4.7	2.26
10	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47
11	Nintendogs	DS	2005	Simulation	Nintendo	9.07	11	1.93	2.75
12	Mario Kart DS	DS	2005	Racing	Nintendo	9.81	7.57	4.13	1.92
13	Pokemon Gold/Pokemon Silver	GB	1999	Role-Playing	Nintendo	9	6.18	7.2	0.71
14	Wii Fit	Wii	2007	Sports	Nintendo	8.94	8.03	3.6	2.15
15	Wii Fit Plus	Wii	2009	Sports	Nintendo	9.09	8.59	2.53	1.79
16	Kinect Adventures!	X360	2010	Misc	Microsoft Game Studios	14.97	4.94	0.24	1.67
17	Grand Theft Auto V	PS3	2013	Action	Take-Two Interactive	7.01	9.27	0.97	4.14

Dataset: wb_health_population

Search Metrics & Columns

Metrics (Showing 6 of 6)

- sum_SP_POP_TOTL
- sum_SH_DYN_AIDS
- COUNT(*)
- sum_SP_RUR_TOTL_ZS
- sum_SP_DYN_LE00_IN
- sum_SP_RUR_TOTL

Columns (Showing 50 of 328)

- region
- year
- NY_GNP_PCAP_CD
- SE_ADT_1524_LT_FM_ZS
- SE_ADT_1524_LT_MA_ZS
- SE_ADT_1524_LT_ZS
- SE_ADT_LITR_FE_ZS
- SE_ADT_LITR_MA_ZS
- SE_ADT_LITR_ZS
- SE_ENR_ORPH
- SE_PRM_CMPT_FE_ZS
- SE_PRM_CMPT_MA_ZS
- SE_PRM_CMPT_ZS
- SE_PRM_ENRR
- SE_PRM_ENRR_FE
- SE_PRM_ENRR_MA
- SE_PRM_NENR
- SE_PRM_NENR_FE
- SE_PRM_NENR_MA
- SE_SEC_ENRR
- SE_SEC_ENRR_FE
- SE_SEC_ENRR_MA

RUN SAVE

TIME COLUMN: year

TIME RANGE: 2011-01-01 ≤ col < 2011-01-02

Query

SERIES: region ENTITY: country_name

X AXIS: f(x) sum_SP_RUR_TOTL_ZS

Y AXIS: f(x) sum_SP_DYN_LE00_IN

FILTERS: country_code NOT IN ('TCA', 'MNP', ...)

BUBBLE SIZE: f(x) sum_SP_POP_TOTL

MAX BUBBLE SIZE: 50

SERIES LIMIT: 0

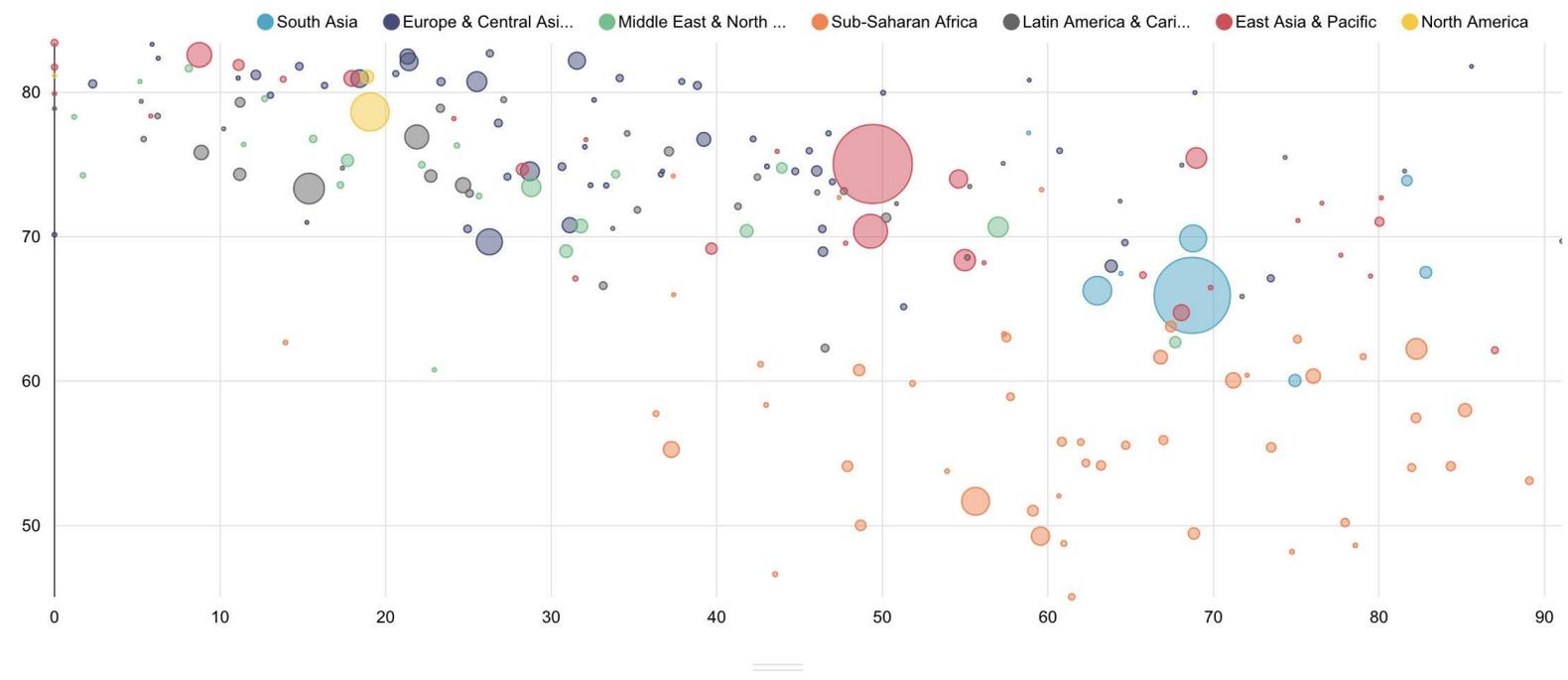
Chart Options

COLOR SCHEME: [Color palette]

LEGEND:

Life Expectancy VS Rural % Altered

201 rows 00:00:00.28



Data

VIEW RESULTS VIEW SAMPLES 201 rows retrieved

region	year	NY_GNP_PCAP_CD	SE_ADT_1524_LT_FM_ZS	SE_ADT_1524_LT_MA_ZS	SE_ADT_1524_LT_ZS	SE_ADT_LITR_FE_ZS	SE_ADT_LITR_MA_ZS	SE_ADT_LITR_ZS
South Asia	1293840000000	570	0.5189674497	61.879070282	46.9900512695	17.6120624542		4
Europe & Central Asia	1293840000000	4390	1.0012649298	98.7313537598	98.7911911011	95.6914825439		9
Middle East & North Africa	1293840000000	4580	null	null	null	null		ni
North America	1293840000000	107530	null	null	null	null		ni
Sub-Saharan Africa	1293840000000	3690	null	null	null	null		ni

« 1 2 3 4 5 »

Типичные требования для разработчика Superset:

- Опыт работы от 2 лет в роли BI-разработчика в Apache Superset;
- Опыт работы в других BI-системах (Power BI, Tableau или Qlik);
- Опыт проектирования и построения витрин, осуществление контроля качества данных;
- Умение писать сложные SQL-запросы (все типы соединений, подзапросы, аналитические функции) и оптимизировать их;
- Знание Python на уровне самостоятельного извлечения данных из различных источников и использования / разработки библиотек для построения сложных визуализаций;
- Опыт работы с ClickHouse, Greenplum, PostgreSQL, MS SQL или с другими СУБД.

Руководство по разработке отчетов в SDP BI

<https://developers.sber.ru/docs/ru/sdp/sdpanalytics/guidelines-reports-SDPBI>

Учебный курс BI-платформа RT.DataVision

https://www.youtube.com/playlist?list=PLAy-0-KaZdl1KiVL64Bd7n5u2Z48_5rQg

	Apache Superset	Русский BI	Qlik Sense
Проект по внедрению отчетности по блокам Производство, Продажи, Логистика			
Источники	1C, PSQL/MsSQL, excel	1C, PSQL/MsSQL, excel	1C, PSQL/MsSQL, excel
DWH	Нужен	Нужен	Нет
1. Техническое задание, ч/ч	600	500	500
2. Установка и настройка ПО (сервер) + BIView, ч/ч	150	80	50
в OS - Devops, коннекторы между компонентами			
3. Разработка модели данных или DWH, ч/ч	1500	1500	650
4. Разработка интерфейса, ч/ч	500	500	400
5. Вывод решения в среду тестирования, ч/ч	250	200	150
6. Вывод решения в промышленную эксплуатацию, ч/ч	200	150	100
Итого человеко-часов	3200	2930	1850
Итого человеко-дней	400	366	231
Стоимость 1 человеко-часа, без НДС, руб	4 200,00 ₽	4 200,00 ₽	4 200,00 ₽
ИТОГО СТОИМОСТЬ ПРОЕКТА без НДС, руб	13 440 000,00 ₽	12 306 000,00 ₽	7 770 000,00 ₽
ИТОГО СТОИМОСТЬ ПРОЕКТА с НДС, руб	16 128 000,00 ₽	14 767 200,00 ₽	9 324 000,00 ₽
Стоимость лицензий на ПО и лицензионной поддержки, 500 пользователей			
Лицензии 1 ый год	0,00 ₽	9 900 000,00 ₽	22 410 000,00 ₽
ТП 1 ый год	0,00 ₽	1 980 000,00 ₽	вкл. в стоимость лицензий
Лицензии 2 ой год	0,00 ₽	0,00 ₽	22 410 000,00 ₽
ТП 2 ой год	0,00 ₽	1 980 000,00 ₽	вкл. в стоимость лицензий
Лицензии 3 ий год	0,00 ₽	0,00 ₽	22 410 000,00 ₽
ТП 3 ий год	0,00 ₽	1 980 000,00 ₽	вкл. в стоимость лицензий
Специалисты в штат для поддержки и развития решения с hh.ru			
SQL разработчик, 1 человек	150 000,00 ₽	150 000,00 ₽	0,00 ₽
Дата-инженер, 1 человек	100 000,00 ₽	200 000,00 ₽	0,00 ₽
Разработчик BI, 1 человек	100 000,00 ₽	100 000,00 ₽	150 000,00 ₽
Итого в год	4 200 000,00 ₽	5 400 000,00 ₽	1 800 000,00 ₽

Стоимость владения

1-ый год			
Совокупная стоимость владения:	20 328 000,00 ₽	32 047 200,00 ₽	33 534 000,00 ₽
Средняя совокупная стоимость на пользователя:	40 656,00 ₽	64 094,40 ₽	67 068,00 ₽
2-ой год			
Совокупная стоимость владения:	24 528 000,00 ₽	39 427 200,00 ₽	57 744 000,00 ₽
Средняя совокупная стоимость на пользователя:	49 056,00 ₽	78 854,40 ₽	115 488,00 ₽
3-й год			
Совокупная стоимость владения:	28 728 000,00 ₽	46 807 200,00 ₽	81 954 000,00 ₽
Средняя совокупная стоимость на пользователя:	57 456,00 ₽	93 614,40 ₽	163 908,00 ₽
4-ый год			
Совокупная стоимость владения:	32 928 000,00 ₽	54 187 200,00 ₽	106 164 000,00 ₽
Средняя совокупная стоимость на пользователя:	65 856,00 ₽	108 374,40 ₽	212 328,00 ₽
5-ый год			
Совокупная стоимость владения:	37 128 000,00 ₽	61 567 200,00 ₽	130 374 000,00 ₽
Средняя совокупная стоимость на пользователя:	74 256,00 ₽	123 134,40 ₽	260 748,00 ₽

Кейс 1: Производитель мебели

Сильные стороны внедрения Open-source BI:

- Open-source BI обычно требуют меньших первоначальных затрат.

Слабые стороны внедрения Open-source BI:

- Небольшой ИТ-отдел компании может не справиться с внедрением и текущим обслуживанием Open-source BI
- Open-source BI могут потребовать значительного времени на внедрение и обучение сотрудников эффективному использованию системы.
- Компания может столкнуться с проблемами интеграции с существующей системой 1С.

План внедрения и команда специалистов:

- Соберите специальную команду по внедрению BI, включая таких специалистов, как менеджер проекта, архитектор данных, разработчик BI и аналитик. Команда будет поддерживать связь с ИТ-отделом и другими заинтересованными сторонами.
- Начните с пилотного проекта, сфокусированного на конкретном отделе или типе отчета.
- Используйте уроки пилотного проекта для распространения BI-системы на другие области.
- Проводите регулярные тренинги, чтобы сотрудники могли эффективно использовать новую систему.

Кейс 2: B2B рынок строительных материалов и услуг

Сильные стороны внедрения Open-source BI:

- Экономически более выгодно (меньше и вначале и не нужно покупать лицензии на потребителей), можно превратить в статью заработка
- Лучшие возможности по доработкам (можно написать что угодно)
- Возможность программировать внутри своего решения (лучше интеграция)

Слабые стороны внедрения Open-source BI:

- Более сложная поддержка и развертывание
- Хорошая квалификация разработчиков поможет с этим справиться

План внедрения и команда специалистов:

- Выделите выделенные ресурсы из существующей команды разработчиков для проекта внедрения BI.
- Начните с пилотного портала для небольшой группы поставщиков и потребителей, постепенно распространяя его на всех пользователей.
- Регулярно собирайте отзывы пользователей и итеративно дорабатывайте дизайн и функциональность портала.

Кейс 3: Банк

Сильные стороны внедрения Open-source BI:

- Решение полностью под контролем
- Не зависим от санкций
- Дешевле
- Возможности использования инноваций (ML, DL, AI) и других разработок комьюнити

Слабые стороны внедрения Open-source BI:

- Нет ФСТЭК, реестр Минцифр
- Банки часто требуют высокого уровня безопасности данных, а Open-source BI могут не соответствовать этим стандартам.

План внедрения и команда специалистов:

- Команда
- Поэтапное внедрение
- Обучение и развитие компетенций: Создайте центр компетенции BI для обеспечения постоянного обучения и поддержки. Это будет способствовать развитию культуры, основанной на данных, и обеспечит эффективное использование новой системы сотрудниками.
- Интеграция и непрерывность



Sergey Gromov

mob. +7 (921) 942-33-75 / skype: sgromych

LinkedIn: <http://ru.linkedin.com/in/gromovsergey>

e-mail: grom@biconsult.ru