

デジタルマーケティングで 活用する BigQuery 入門

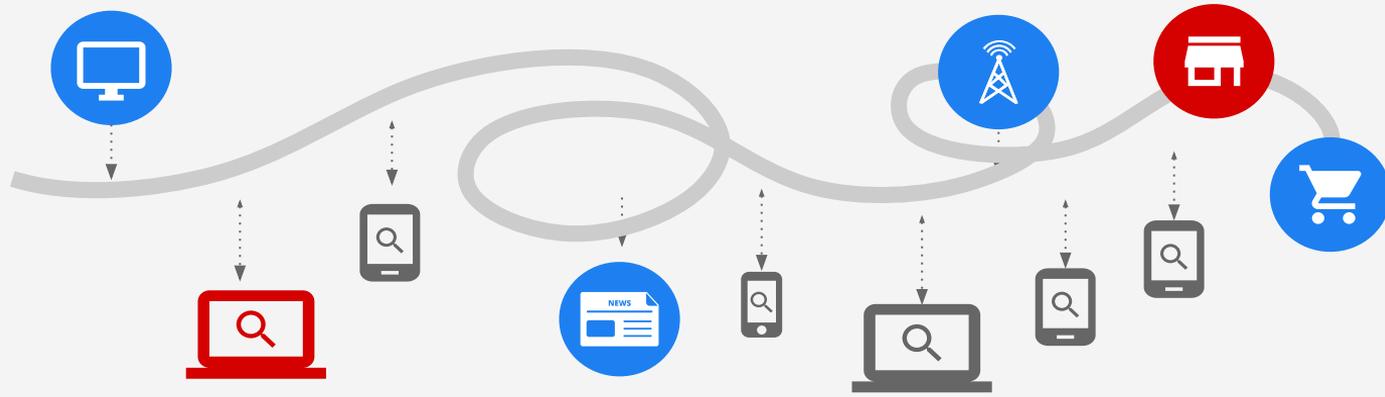
Google Cloud
カスタマーエンジニア
千川 浩平

マーケティングデータ基盤 (CDP)が 求められる背景とその機能	01
Google の提供するサービスで実現する CDP	02
ユースケース紹介	03
まとめ	04

01

マーケティング データ基盤 (CDP) が求められる背景とその機能

カスタマージャーニーは年々複雑に



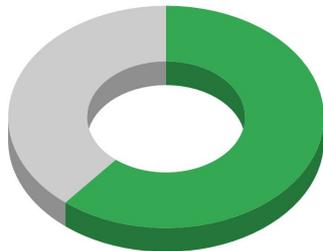
ビジネスの成功には

あらゆるタッチポイントで
顧客の興味、関心に寄り添うことが不可欠

61%



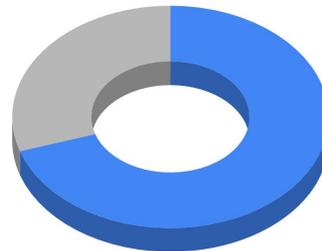
自分の要望に基づく
対応を企業に
期待している人の割合



63%



購入履歴に基づいて
パーソナライズされた
エクスペリエンスを
期待している人の割合



社内のおちこちに分散した顧客データ

全体像の把握は困難

13 %

社内の顧客データを
分析に活用できている
と答えた企業の割合

Forbes Insights, 2018

26 %

自社のマーケティング
分析ツールはデータの
統合レベルが高いと思う
マーケティング担当者の割合

Forrester, 2016



マーケティングに活用できるデータの例



マーケティング データ基盤(CDP)に求められる要素



データ収集・管理

以下のような種々のソースからデータを収集する

- Web サイト(タグ経由)
- CRM データベース
- コールセンターログ
- 売上データ(オンライン・オフライン)
- 広告や媒体データ



ID 統合

異なるデータソースの顧客情報を統合し、同一ユーザとして分析できるようにする



複雑な分析

機械学習(セグメンテーション・購入確率の計算・LTV 予測など)と柔軟なレポート機能

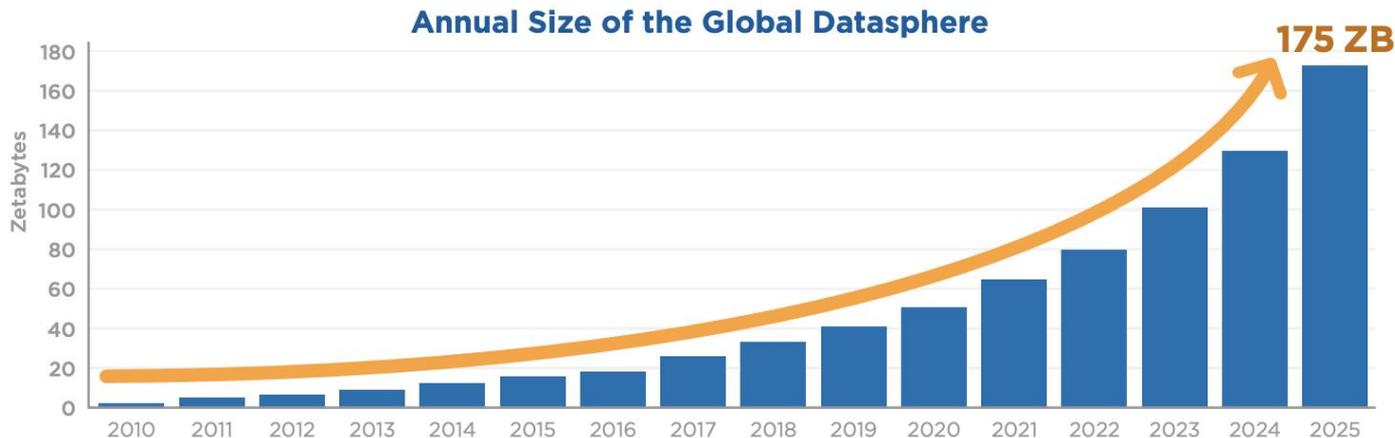


施策実施

複数チャネル横断のパーソナライズされた接点によるマーケティング施策の実施

世界全体のデータ量は 2018 年の 33 ZB から 2025 年には 175 ZB へと増加する見通し

IDC*



*<https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>, Nov 2018

02

Google の提供するサービスで 実現する CDP

すでにお使いの製品を活用して CDP の機能・価値を実現する

Google Cloud



 Google Marketing Platform

Google は各機能に必要なソリューションを提供しています



データ収集・管理

既存のタグを使用

メディアとサイトデータの
連携をクリックだけで実現

オフラインデータの
ストリーミング取り込み

各種 API



BigQuery

Data Transfer Service



Pub/Sub



ID 統合

キー・ID で Join するため
のクラウドコンポーネント群



Data Fusion



DataFlow



BigTable



BigQuery



複雑な分析

事前構築済のモデル

各種の ML オプション



BigQuery ML



Vertex AI



AutoML Tables



TensorFlow



施策実施

ダッシュボード

広告・サイト分析との
インテグレーション



Data Studio



Analytics 360

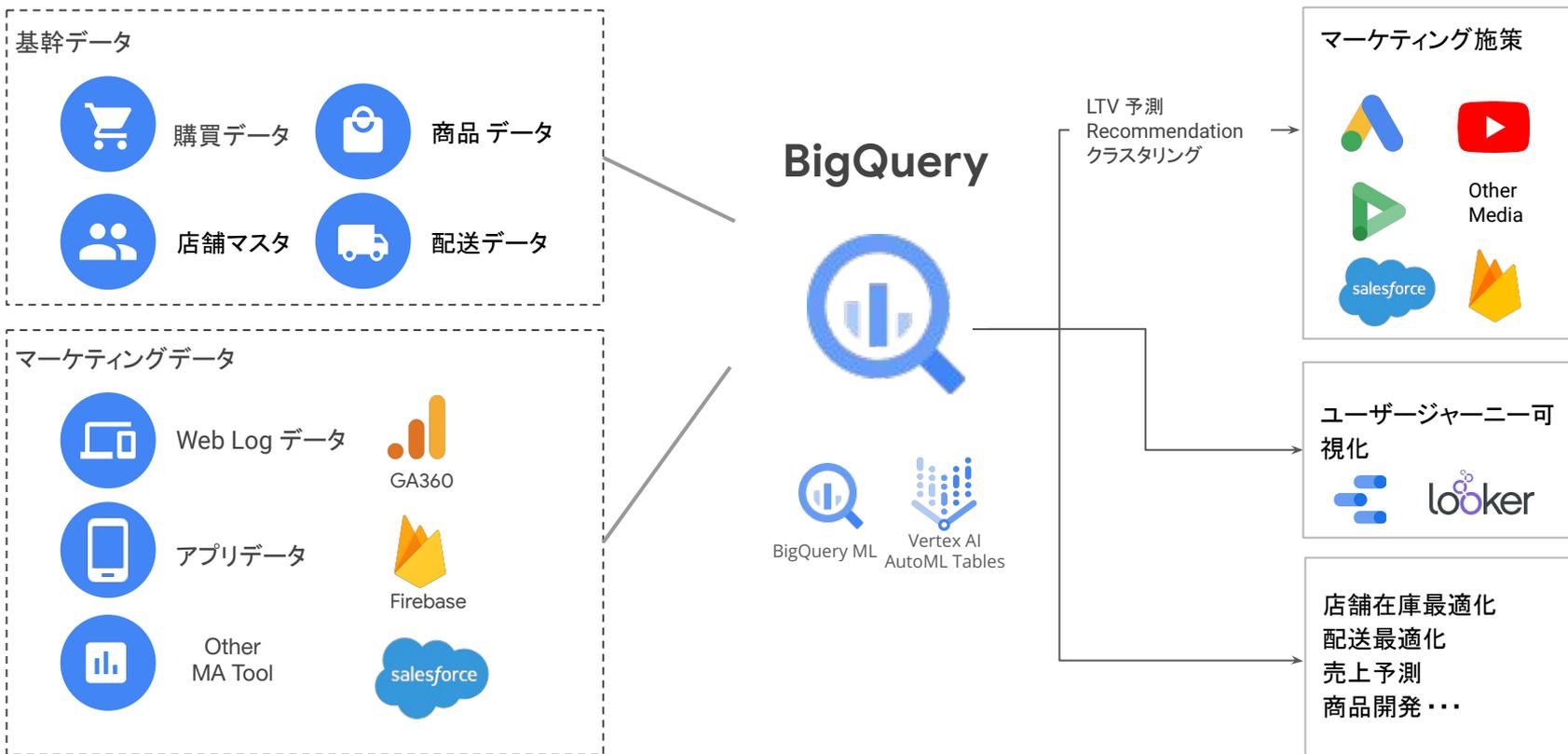


Campaign Manager



Ads Data Hub

Google Cloud で実現するマーケティングデータ活用基盤



データ収集・加工

分析・予測

アクション
顧客体験

既存 CDP との比較

既存の CDP 製品

- ・ワンパッケージ
- ・長いリードタイム(導入・データ蓄積)
- ・データの拡張性に制限あり

Google の提供する機能で実現する CDP

- ・必要なコンポーネントを取捨選択
- ・短いリードタイム(既存資産の活用)
- ・**大量データに対する柔軟性が強み**

Google Cloud では: ネイティブアプリ、Web 解析連携

/ ワンクリックで BigQuery にデータをストリーミング連携、他のデータとの統合分析が可能

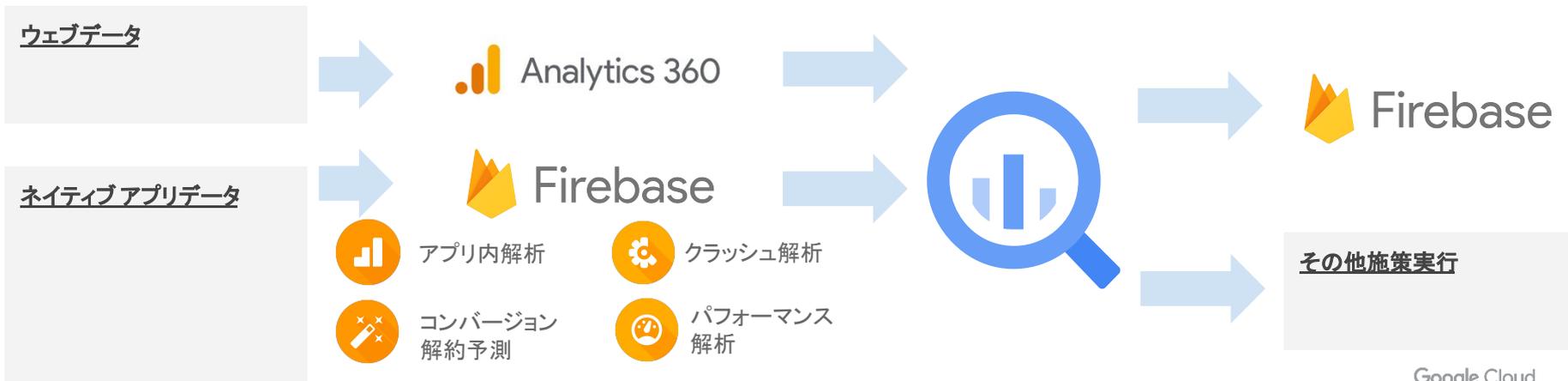
/ **Web**: Google Analytics 360, Google Analytics 4 プロパティ

/ **ネイティブ**: Google Analytics for Firebase, その他 Firebase サービス

GA / GA4F の raw data を export できるのは BigQuery のみ

/ Firebase A/B Testing, Cloud Messaging (プッシュ) などが BigQuery セグメントの定期インポートをサポート

/ 「打ち手」の実行までを自動化、マーケターでもできるように



ID 統合のイメージ

- それぞれのテーブルにキーをもたせ、それらを join
- どのプラットフォームでも SQL による実装が必要で、**優劣なし**

 Web Analytics Data

datetime	ga_id	login_uid	page	...
2018/1/1 12:00:23	ewrgwrw_rr	1234XX	/special/campX	
2018/1/2 12:00:23	grtcdfv_dw	3456YY	/mobile	
2018/1/3 12:00:23	ferfregyj_jh		/television	

 Product Purchase Data

datetime	login_uid	email_address	product	...
2018/1/1 15:12:03	1234XX	aaa@gmail.com	prod1	
2018/1/2 19:40:20	1234XX	aaa@gmail.com	prod2	
2018/1/1 12:00:23	3456YY	bbb@hotmail.com	prod3	

 Smartphone App Usage

datetime	login_uid	device_id	App Name	...
2018/1/1 15:12:03	1234XX	55DSSGVH	App 1	
2018/1/2 19:40:20	1234XX	55DSSGVH	App2	
2018/1/1 12:00:23	55GHYU	6HJFDF68	App1	

 Unified Customer Profile

email	ga_id	saw_campaign_X	web_vis_frequency	login_uid	has_prod1	has_prod2	...	use_App1	use_App2
aaa@gmail.com	ewrgwrw_rr	Y	3.4	1234XX	Y	Y		Y	N
bbb@hotmail.com	grtcdfv_dw	N	5.1	3456YY	N	Y		N	N
	ferfregyj_jh		1.2					N	N
				55GHYU				N	Y

大量のマーケティングデータを処理できる BigQuery

エンタープライズ向け
分析用データ ウェアハウス

エクサバイト規模のストレージと
ペタバイト規模の SQL クエリ

セキュア、耐久性
メンテナンス フリー



フルマネージドでサーバレス

Unique

ストリーミング データの
リアルタイム分析

Unique

ML と GIS をビルトイン

Unique

ハイスピードでインメモリの
BI エンジン

Unique

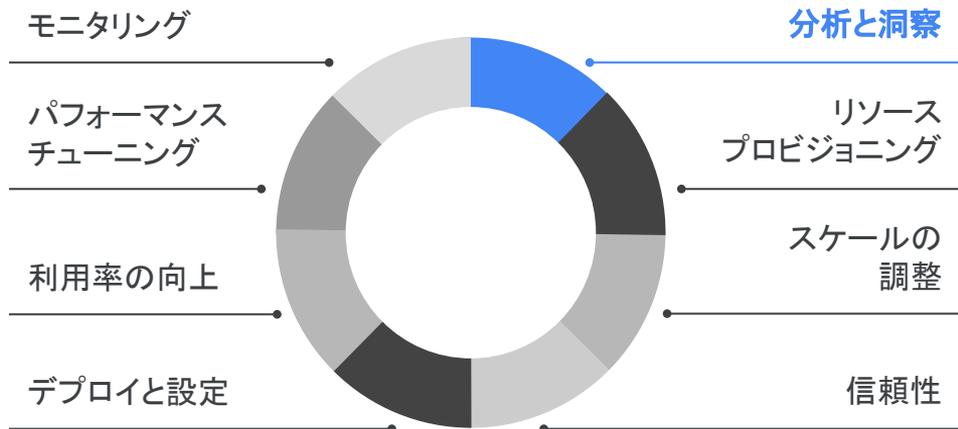
大量のマーケティングデータを処理できる BigQuery





BigQuery | サーバーレス データ分析

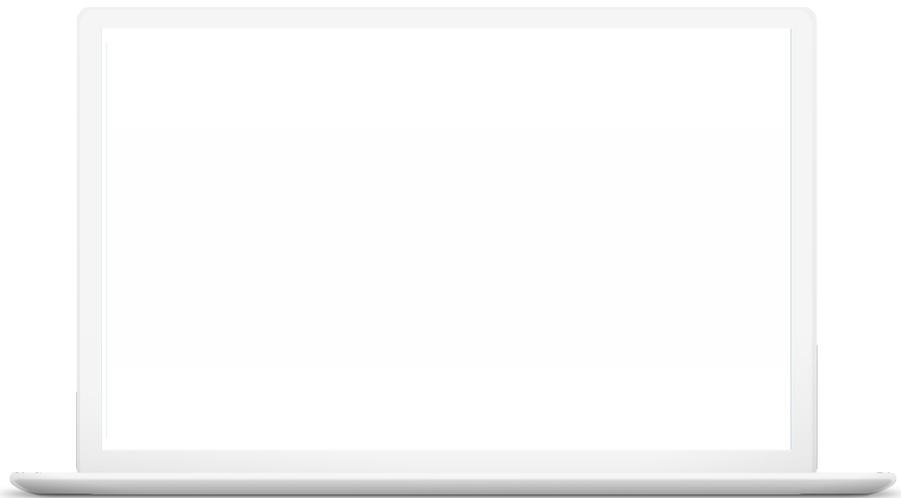
従来のデータウェアハウスとの違い



BigQuery のサーバーレス分析



BigQuery ML を使用して AI の基盤を構築する



- 回帰
- 顧客セグメンテーション: K-means クラスタリング
- 製品のレコメンなど: 行列分解
- Tensorflow モデルのインポートし BigQuery で予測
- BigQuery で Tensorflow DNN モデルを作成する

1

BigQuery からデータを移動せずに ML インシニアチブを
実行

2

開発速度を上げるために
BigQuery で SQL のモデルを
繰り返す

3

一般的な ML タスクとハイ
パー パラメーター調整を
自動化

ドキュメント: [BigQuery ML の概要](#)

BigQuery パフォーマンスの進化 (1 PB のテーブルをスキャン)

Query ?

```
SELECT *  
FROM [retail_pe  
WHERE salesOrder
```

Run Query Save Query Save

complete (245.7s elapsed, 1.09 PB processed)

salesOrderKey
1104796800000-195488
1104796800000-195488

2016 年
245.7 秒

CAN-FILTER */

```
CT *  
OM google.com:bigquery-peta  
E customerKey = "14408064000
```

Processing location: US No cached results

Run query Save query

Query results SAVE AS

complete (min 53.616 sec elapsed, 402.5 MB processed)

Information Results JSON Export

Some repeated values are hidden to improve performance

2018 年
113 秒

```
1 SELECT  
2 *  
3 FROM  
4 [retail_petabyte.sales_  
5 WHERE  
6 customerKey = "14408064000
```

No cached results

Run Save query

Query results

Query complete (4.2 sec elapsed, 402.5 MB processed)

Job information Results JSON

Some repeated values are hidden to improve performance

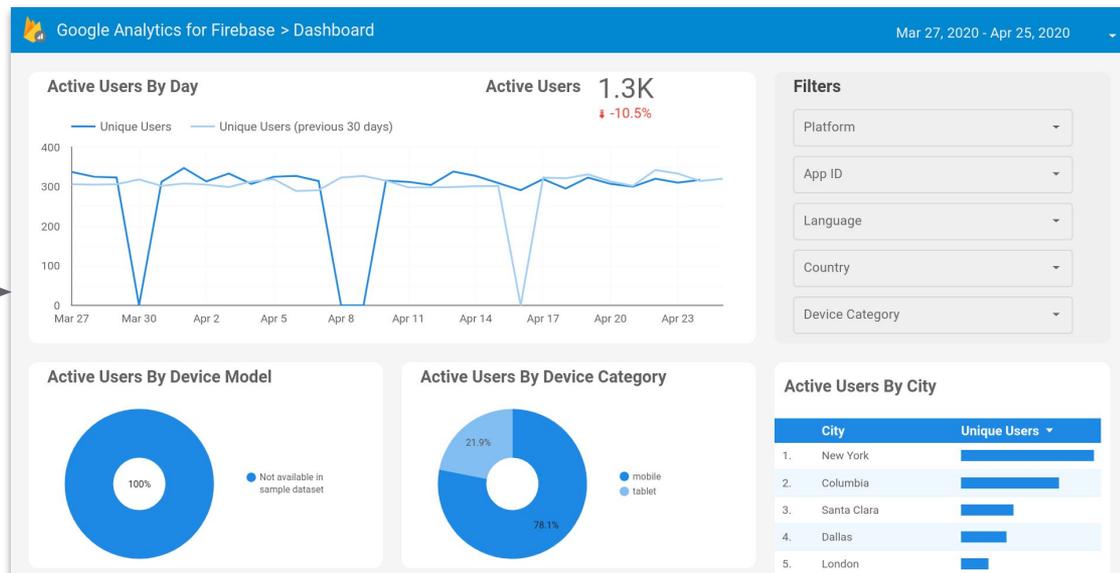
2019 年
4.2 秒

03

ユースケース紹介

1. KPI のモニタリング

- Sheets や Data Portal、Looker で KPI を表示するレポートを作成してモニタリング
 - DAU, WAU, MAU
 - 売上、課金率
 - 継続率、離脱率
 - ect.

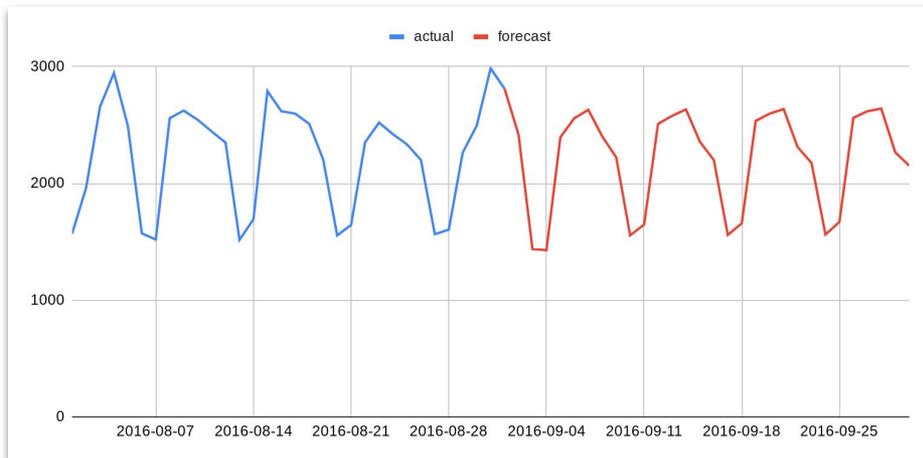


2. KPI の予測

- DAU や売上など、過去の値の推移をもとに、将来の値の推移を予測
- BigQuery ML で時系列予測のモデルを構築して予測

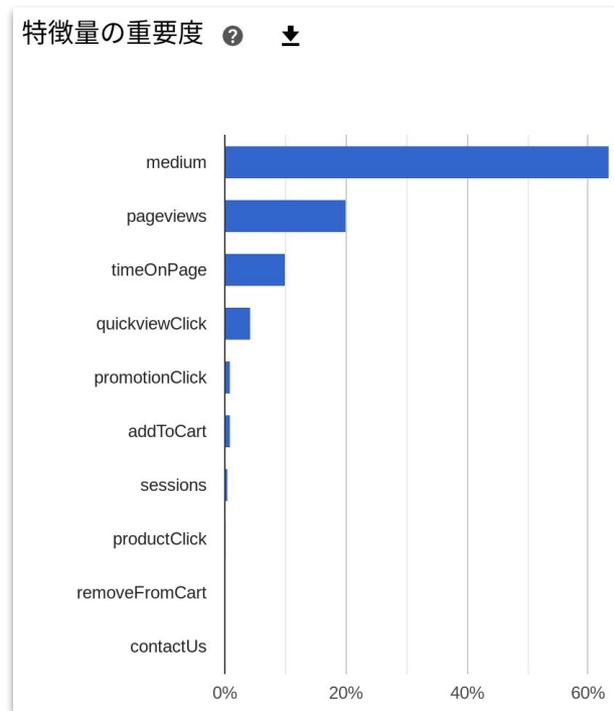
```
1 create or replace model `gcj-gaming-analytics.ga.predict_dau`  
2 options(model_type='arima', time_series_data_col='dau', time_series_timestamp_col='date') as  
3 select  
4   timestamp(date) as date  
5   , count(distinct fullVisitorId) as dau  
6 from  
7   `gcj-gaming-analytics.ga.sessions`  
8 where  
9   date between '2016-08-01' and '2016-08-31'  
10 group by  
11   date  
12 order by  
13   date
```

```
1 select  
2   date(forecast_timestamp) as date  
3   , forecast_value as forecast  
4 from  
5   ml.forecast(model `gcj-gaming-analytics.ga.predict_dau`, struct(30 as horizon))  
6 order by  
7   date
```



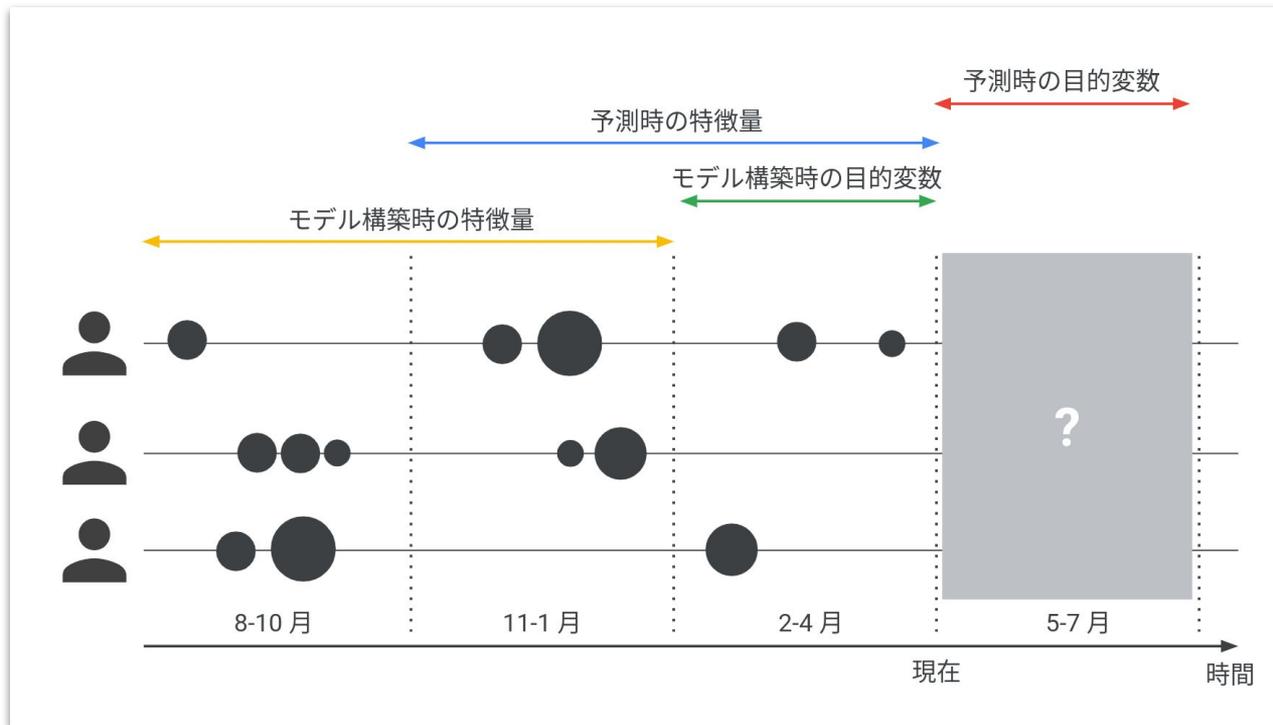
3. KPI 改善施策の検討

- AutoML Tables の「特徴量の重要度」を表示する機能を使用
- 「KPI に大きな影響を与える要因として、どのようなものが考えられるか」を調査し、それをもとにKPI 改善施策を検討して実施
 - 継続率: N 日後に継続するユーザとそうでないユーザの違い
 - 課金率: 課金するユーザとそうでないユーザの違い



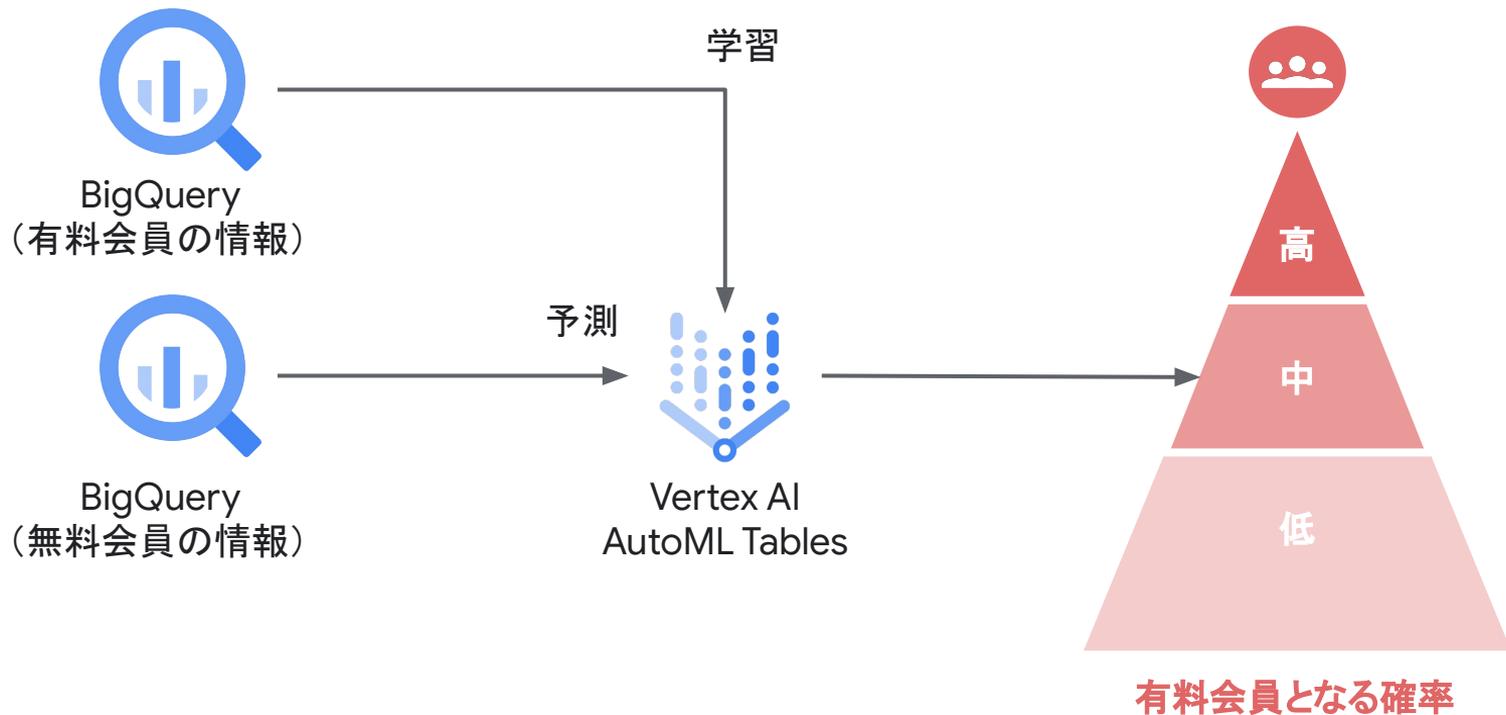
4. LTV(顧客生涯価値)の予測

- 過去のユーザの購買履歴をもとに、各ユーザのLTV(顧客生涯価値)を予測
- RFM(最新購入時期、購入頻度、購入金額)などを特徴量として使用



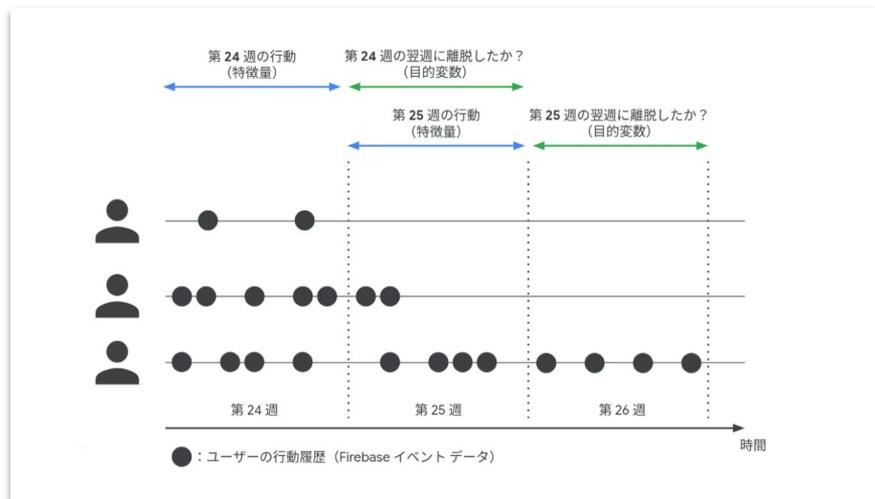
5. 有料会員の予測

- 既存の有料会員の行動履歴やユーザ属性をもとに、既存の無料会員のうち、将来的に有料会員となりそうなユーザを予測

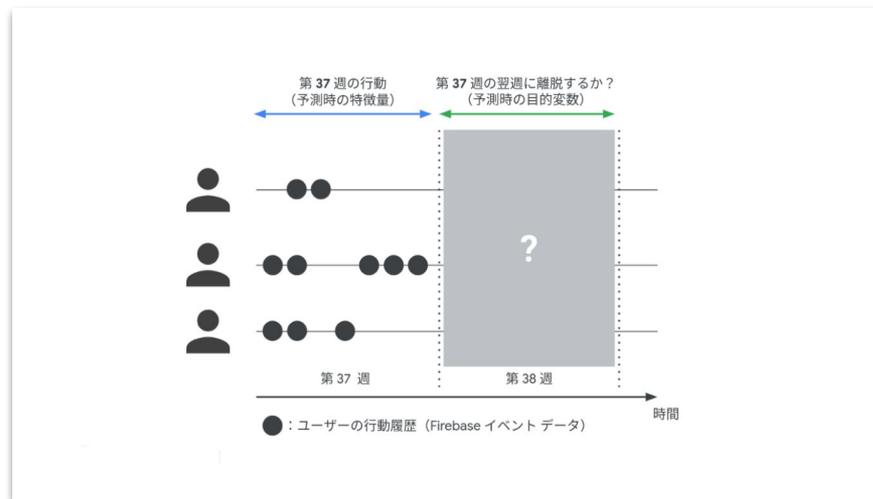


6. 離脱ユーザの予測

- 過去のユーザの行動履歴をもとに、将来的な離脱ユーザを予測
- AutoML Tables の「特徴量の重要度」の表示機能を利用して、「どのような要因が離脱に大きな影響を与えていそうであるか」を確認することが可能



予測モデルの構築



予測

7. コンテンツのレコメンド

- 各ユーザの購買履歴や視聴履歴をもとに、各ユーザに好みに近いユーザが購入または視聴したコンテンツをレコメンド
- BigQuery ML のレコメンド モデルを使用

	コンテンツ A	コンテンツ B	コンテンツ C	コンテンツ D	レコメンド
ユーザ 1	✓				コンテンツ C
ユーザ 2		✓		✓	コンテンツ C
ユーザ 3	✓		✓		コンテンツ D
ユーザ 4			✓	✓	コンテンツ A, B

```
1 create or replace model `gcj-gaming-analytics.ga.recommend_products`  
2 options(model_type='matrix_factorization') as  
3 select * from `gcj-gaming-analytics.ga.user_item_ratings`
```

04

まとめ

まとめ

1. 多様化する顧客との接点のデータを統合してビジネスに変革を
2. Google のプラットフォームで既存資産を活用してクイックに
3. 機械学習・リアルタイム性などより高度な施策に挑戦

データを用いてより良い顧客体験を！！

Thank you