

# 経営変革を実現するモダンデータ分析基盤を解説 - trocco® & BigQuery 構築事例 -

岩田 匠

株式会社primeNumber、Head of Solution Architect



# 株式会社primeNumber

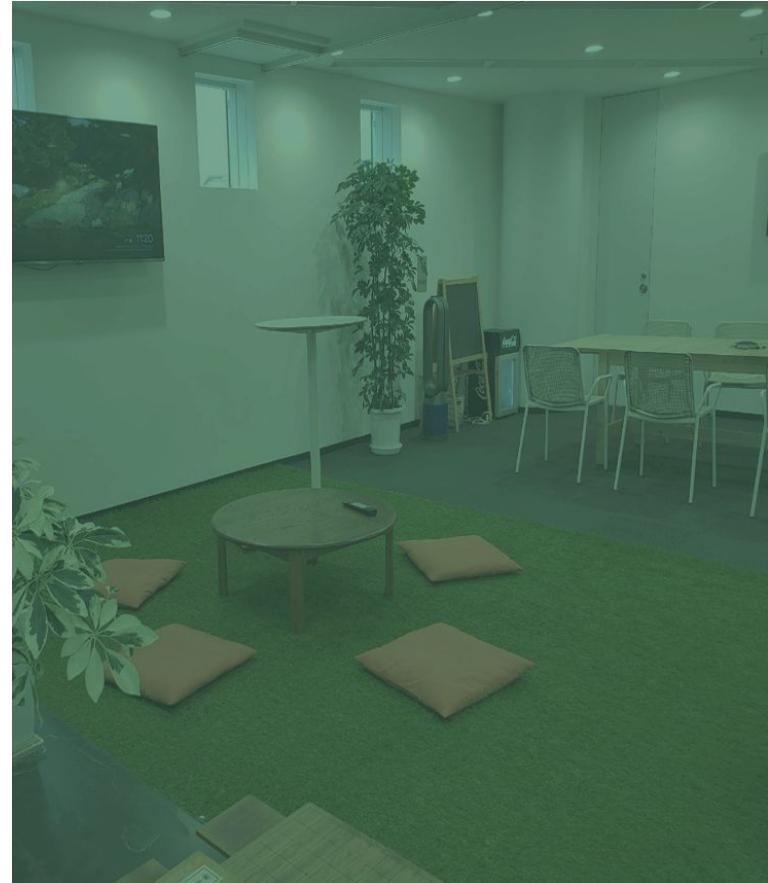
代表取締役CEO  
田邊 雄樹

従業員数  
33名

創業  
2015年11月

2021年12月  
シリーズBラウンド資金調達

シリーズAと合わせて累計14億円以上を調達済み



# あらゆるデータを、 ビジネスの力に変える。

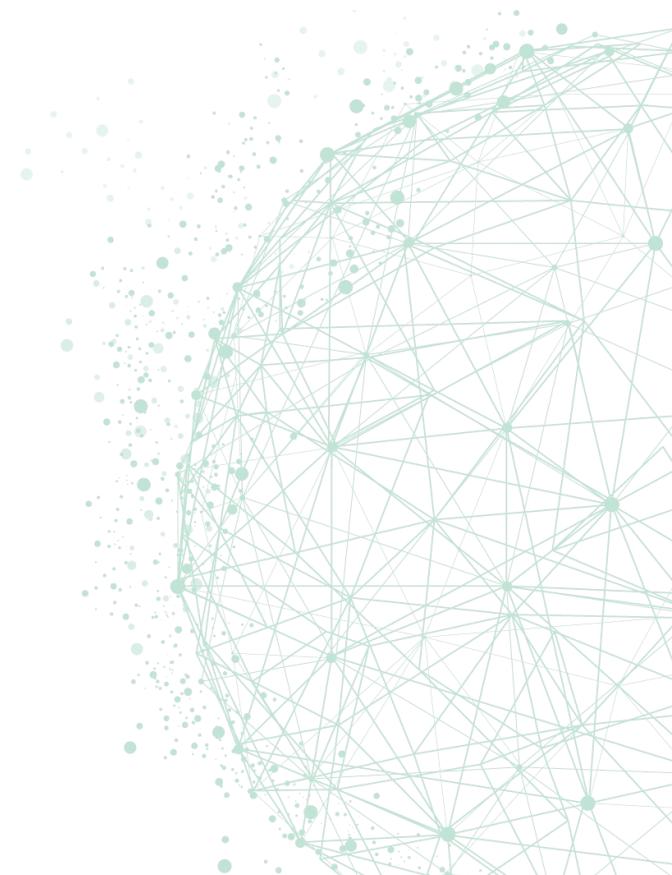
primeNumber は、データテクノロジー カンパニー。

あらゆるデータが爆発的に増えていく時代に、

誰もがすばやく、簡単にデータを使える環境を構築し、

データ活用までのプロセスを最適化。

高度なテクノロジーと独自のアイデアで、世界中のビジネスを支援します。





**primeNumber** は、  
データエンジニアリングにフォーカスを当てた  
3つのサービスを提供しています。

# 経営変革を実現する モダンデータ分析基盤を解説 - trocco® & BigQuery 構築事例 -

金谷 直樹

株式会社サーティキュレーション IT推進室 マネジャー

本日の内容

半年間、経営変革の実現に向けて取り組んだ  
データ分析基盤づくりの学びを共有します。

# 最初に: 本日想定している主な聴講者

- 初めて、データ分析基盤の構築を検討されている方
- 既存のデータパイプラインの保守工数を削減し、エンジニアの時間を創出したい方
- 他社事例に関心をお持ちのデータスチュワード / データエンジニアの皆さん



## 金谷 直樹

株式会社サーキュレーション 経営管理本部  
企画・マーケティング・IT部 IT推進室 マネジャー

インターネット広告企業や研修事業会社の情報システム部門のマネジャーとして、  
ビジネス / コーポレートの各種インフラ基盤構築やシステム導入を通じ、企業のデジタル化を  
推進。2021年5月に株式会社サーキュレーション入社後、経営戦略や事業課題にもとづく会  
社のデジタル戦略を企画・推進する傍ら、エンジニア組織の組成や文化醸成にも注力し、デジ  
タル経営変革を会社全体で巻き起こすために活動中。



どんな課題でも、  
プロがいれば解決できる。



# CIRCULATION

株式会社サーキュレーション

Google Cloud

## OUR VISION

# 世界中の経験・知見が循環する社会の創造

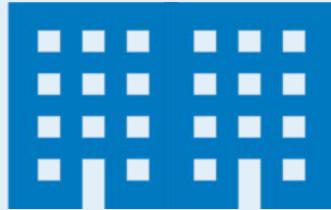
知のめぐりを良くするプラットフォームを提供しています。

# プロシェアリングとは？

## プロをシェアするプラットフォーム

外部プロ人材の経験・知見を複数の企業でシェアし、あらゆる経営課題を解決する

### 法人企業



### プロシェアリングのプロジェクト

#### 経営課題



#### DX,先端技術 課題



#### 新規事業 アイデア創出



#### 事業承継 M&A



### プロ人材



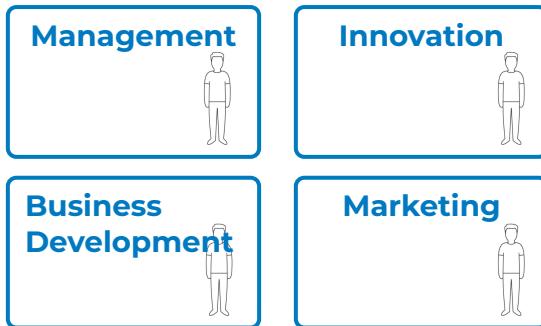
- 専門性で解決すべき経営課題
- 社内の知見や人材の不足

- 週1日から週3日のプロジェクト稼働
  - 一日あたり2時間から6時間の稼働
  - 3ヶ月から12ヶ月のアサイン
- ※アサインするプロ人材や稼働状況に応じて価格は設定

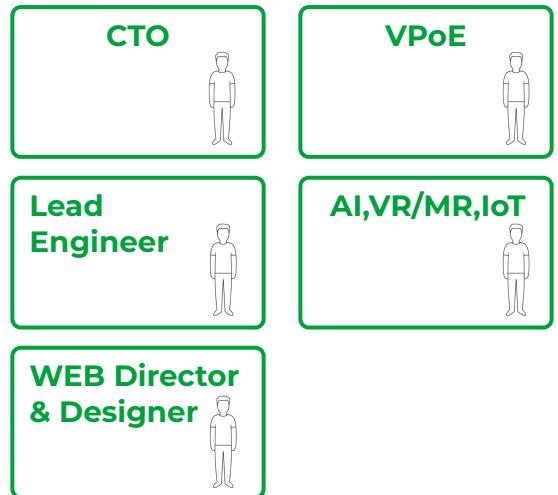
その領域のトップランナー 18,388 名<sup>(1)</sup> のプロ人材リソースを保有



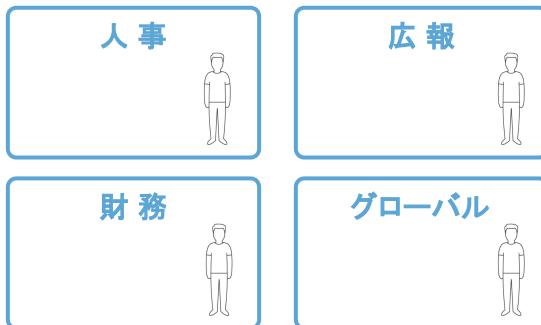
## Business Development - Professional



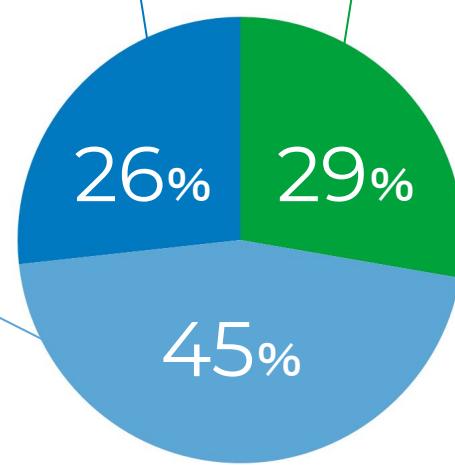
## DX-Professional



## Business Theme - Professional



※その他:M&A、営業等



トッププロが日本で一番集まるプラットフォームへ

スクリーニング

アセスメント

プロジェクト  
評価

注: (1)2021年10月末時点の登録者総数。当社サービス退会済みの方を含めず。



# 01

## データ分析基盤の導入背景、課題

# データ駆動経営に欠かせない3つの構成要素

## 1. データ エンジニアリング

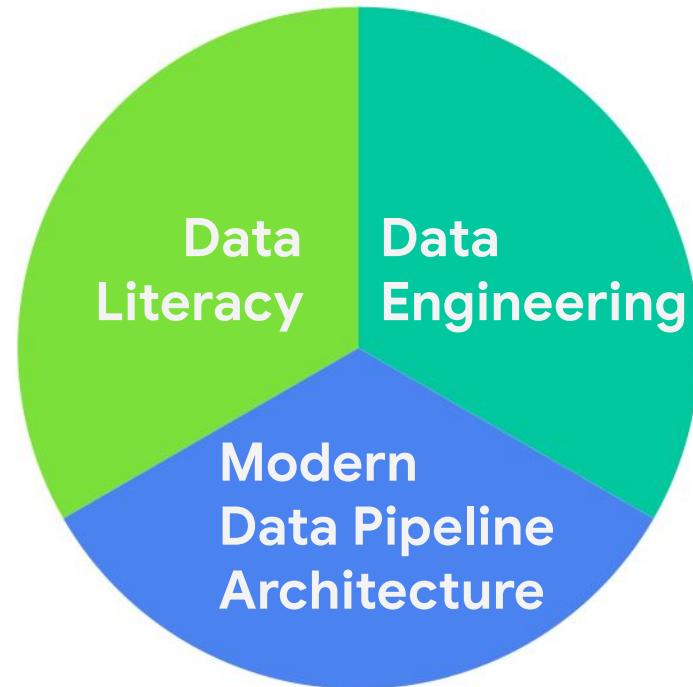
時には業務システムのデータの持ち方まで立ち返り  
ビジネスロジックや構造化データの見直しなど  
深いビジネス理解と多様な実現手腕を磨く

## 2. モダンなデータ分析パイプライン

省力開発・運用が可能なETL / DWH / BIツールの選定が  
日々の高速な仮説検証を支える

## 3. データリテラシーの組織学習

経営者をはじめ意思決定層に対する学習と実践のサイクル  
意思決定や業務の自動化・効率化を導く専任チームの組成



# 半年前の私たちは、全くのゼロからスタート

## 1. データ エンジニアリング

時には業務システムのデータの持ち方まで立ち返り  
ビジネスロジックや構造化データの見直しなど  
深いビジネス理解と多様な実現手腕を磨く

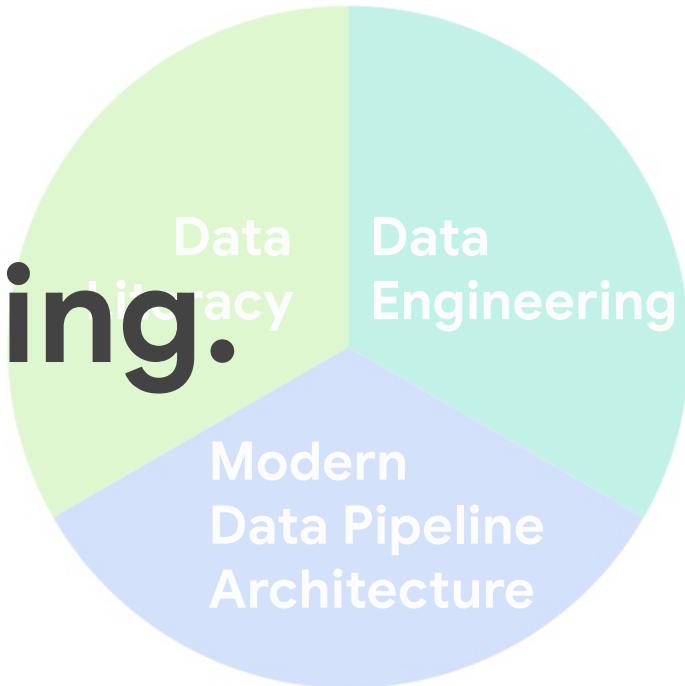
## 2. モダンなデータ分析パイプライン

省力開発・運用が可能なETL / DWH / BIツールの選定が  
日々の高速な仮説検証を支える

## 3. データリテラシーの組織学習

経営者をはじめ意思決定層に対する学習と実践のサイクル  
意思決定や業務の自動化・効率化を導く専任チームの組成

# ...Nothing.



# 経緯 & 背景：組織とビジネスの急成長とともに生じる歪み



社員数  
100→200

- The Model の営業組織体制を形成し、分業型で顧客成功を追求
- × 部門最適への偏りが目立ち始める(増え続ける独自KPI やノウハウの横展開が滞るなど)



CAGR  
72.9%

- 創業来、高い年平均成長率を維持(CAGR 72.9%)
- × 事業成長スピードに対し、経営分析環境の整備やビジネスプロセス改善に対する企画・開発体制が追いつかない



Salesforce  
フル活用

- MA / SFA / CRM では、AIによる予測モデルやレコメンデーション等、イノベーションが加速
- × カスタマイズ領域の保守性やシステム上限・制約に懸念。大掛かりな品質改善計画を着手

# 経緯 & 背景：組織とビジネスの急成長とともに生じる歪み

## コミュニケーション

社員数

100→200

## ビジネスプロセス

○ The Modelの会員組織体制  
を形成し、分業型で顧客成功  
を追求

## 業務システム

○ 部門最適への偏りが目立ち始  
める(増え続ける独自KPI やノ  
ウハウの横展開が滞るなど)



CAGR

72.9%

○ 創業者  
高い年平均成長率を  
維持(CAGR 72.9%)

- × 事業成長スピードに対し、経  
営分析環境の整備やビジネス  
プロセス改善に対する企画・  
開発体制が追いつかない

## 硬直化 の兆し

○ MAによるCRMでのAIによ  
る予測モデルやロボット一  
シテーションによる自動化

- カスタマイズ領域の保守性や
- システム上限・制約に懸念。大  
掛かりな品質改善計画を着手

# 実現したいことの初期イメージ



Salesforce(MA / SFA / CRM)  
会計システム  
人事管理システム  
Google Workspace  
他

事業活動のあらゆるデータを  
分析可能な状態で**蓄積**

正しいデータを**分析**し  
迅速に**意思決定**し  
**価値**に繋げる能力を獲得

# 欲しい機能と要件の具体例

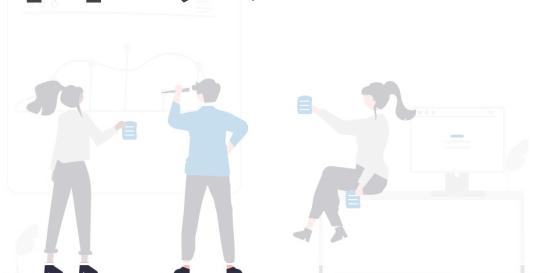
1. 財務分析・原価管理・予実管理
2. リカーリング プロセスの課題発見
3. 主要 KPI の変動要因分析
4. 仮説立案と行動後の迅速な検証
5. 成功モデルの横展開



# プロジェクト ゴール

1. 財務分析・原価管理・予実管理
2. リカーリングプロセスの課題発見  
主導する変動要因の新規発見
4. 仮説立案と行動後の迅速な検証
5. 成功モデルの横展開

# 高速な仮説検証を躊躇なく 繰り返せるようになること



可視性

容易に追跡可能

# しかし、データ活用に関する経験・知見が社内に殆どない

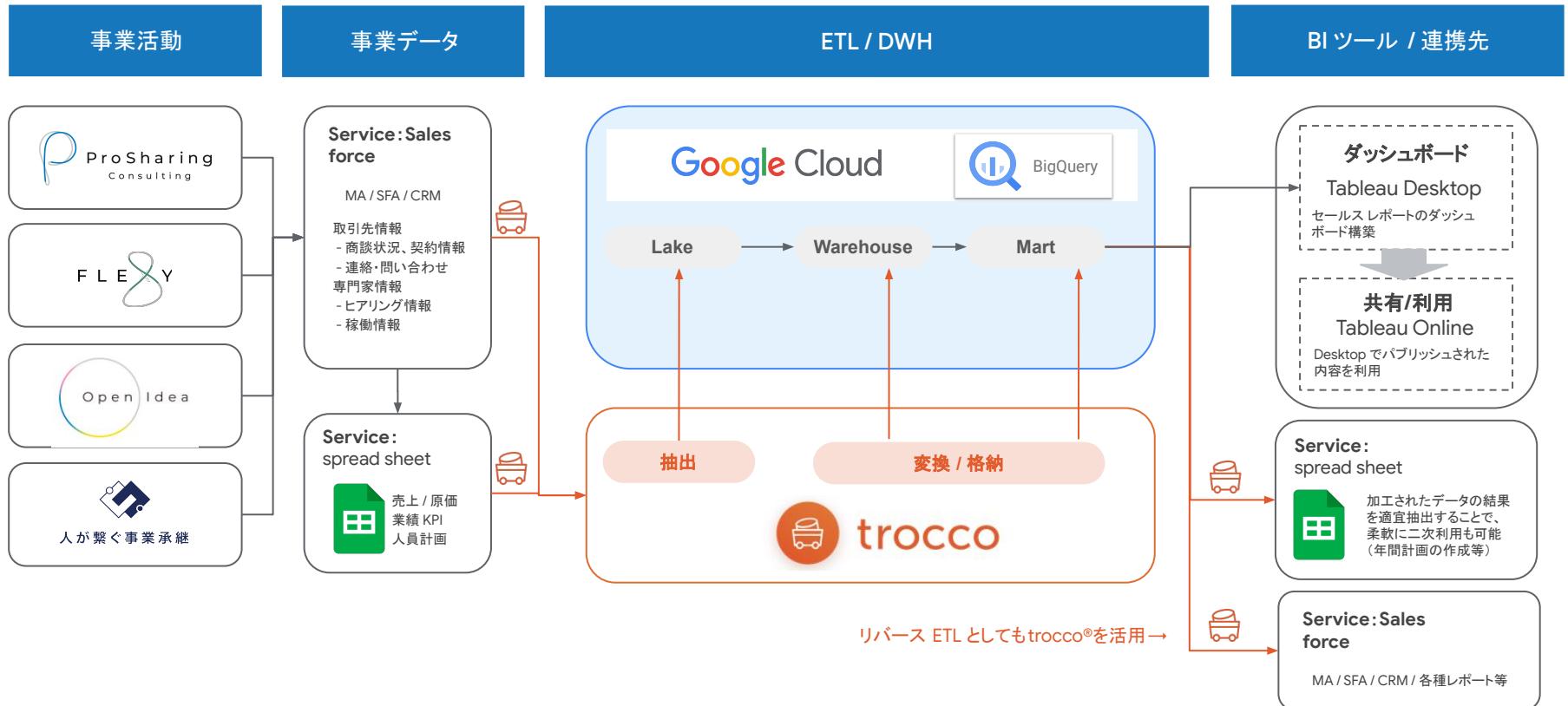
		経営、事業理解	Salesforce Platform アプリ開発、 データベース理解	データ エンジニアリング	データ ビジュアライズ
サービス 提供会社	プロジェクト マネージャー	○	△	×	×
	経営企画、事業企画	○	△	×	×
	コーポレート エンジニア	△	○	×	×
	primeNumber ソリューションサービス	-	○	○	○

primeNumber のソリューションサービス活用でプロジェクトスタート  
この先の内製のデータ組織化を目指し、スキルトランスファーまでを依頼



# アーキテクチャ概要/選定

# データパイプライン概要



重要視:

データエンジニアの時間

インフラ保守 < データ活用

- データパイプラインの **保守に係る工数** を極力減らせること
- よって、クラウド、マネージド サービスをフル活用
- さらに、**エンジニア以外** にも多少触れる余地があること
- ETL には**豊富なコネクタ** が揃っていること
- **学習コスト** が安価であること + 経済コストも考慮

# ETL は一段上の扱いやすさから、trocco® を選定

ETL	DWH	BI
<b>省力開発・運用</b>	○	△～○
<b>エンジニア以外も触れる余地 (簡易な転送設定の追加 / 変更等)</b>	○	△
<b>学習コストの低さ (日本語ドキュメントの充実、サポート体制)</b>	○	△
<b>コネクタ、変換プラグインの充実</b>	○(約 100) Salesforce / Tableau / google に対応しており、必要十分	◎(150 超) Google Cloud 連携の 充実ぶりはさすが
<b>セキュリティ</b>	○	○
<b>費用</b>	△	○
	 trocco	<b>Google Cloud ETL 处理関連</b>
		<b>その他の ETL 处理関連</b>

# DWH は、BigQuery 一択

ETL

DWH

BI

- 迷いなく、「BigQuery」を選択
  - 何より、trocco® と BigQuery の組み合わせにより、省力開発・保守のイメージが湧いたこと
    - trocco®上から、UI によるワークフロー管理や標準 SQL を設定できたり、転送先テーブルが存在しない場合のテーブル自動作成が可能
  - 経済コストに優れ、パフォーマンスへの不安も少ない
  - (母集団が少ない)データエンジニア募集時の保有スキルも意識
- 迷いなくと言いつつ…
  - Snowflake も気になったが、当面のデータ処理量の傾向から生かしきれないかも。(BQ 向き)
  - 個別契約になってしまうことや、コストシミュレーションにやや難しさを感じ、trocco® と BigQuery の連携の高さもあって今回は見送り。(将来検討あり)



# BI ツールはビジュアル機能と簡易さから、Tableau を選定

ETL

DWH

BI

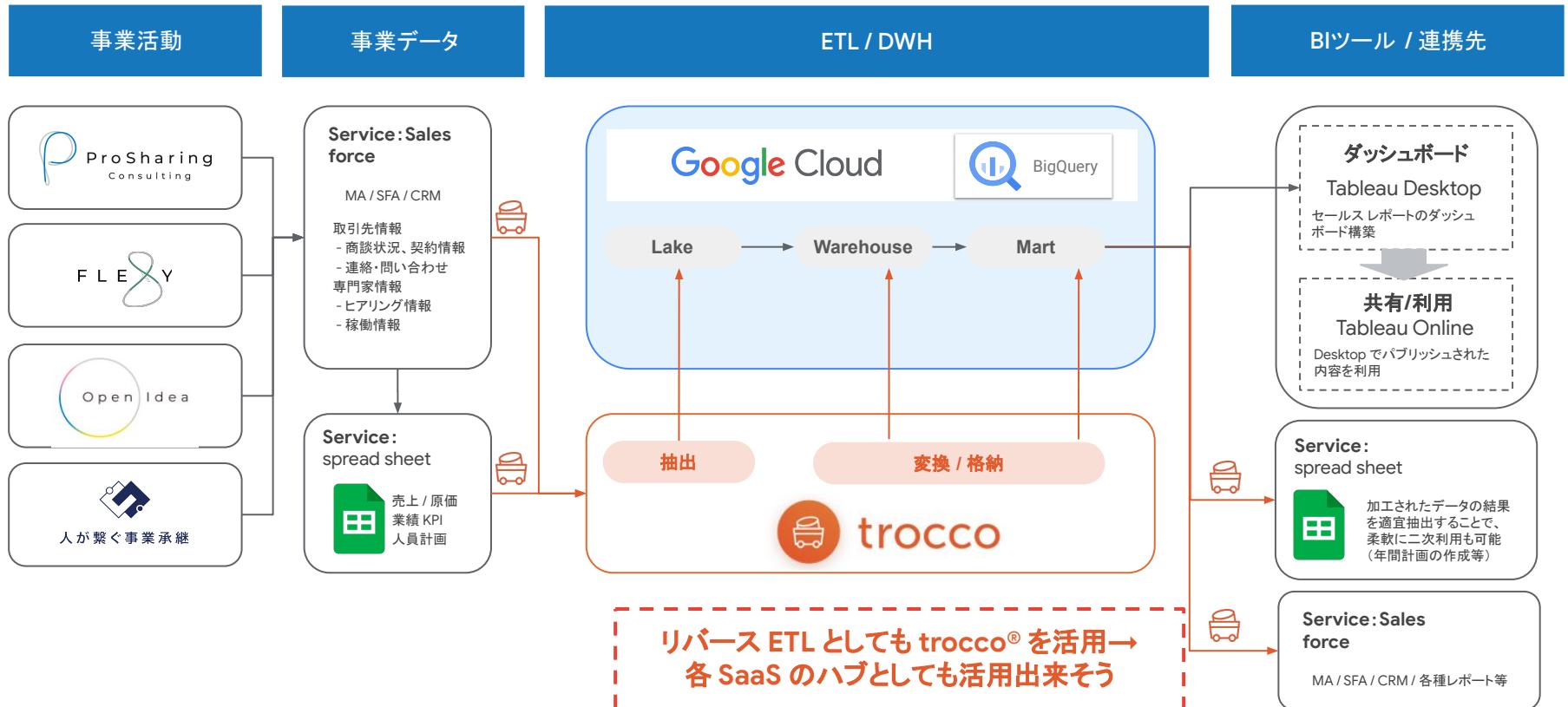
## Tableau

- 標準で強力なビジュアライゼーション
- ノーコードでの分析やローカル利用も可能
- 探索的なデータ分析時に扱いやすいUI
- 初期コストが安価で日本語ドキュメントやサポート、コミュニティも充実



- 一定のデータリテラシーが必要と認識
- LookML によるモデル定義を経由するメリットや Git によるバージョン管理など関心あり
- データ組織やデータの民主化が成熟したときに検討したい

# データパイプライン概要(補足:リバースETL)





# trocco® & BigQuery

## データ分析基盤 導入成果

# データ分析基盤の初期構築と一次リリースまで完了

## 1. 財務分析・原価管理・予実管理 ✓

- a. 80-90 h / 月の工数削減効果(約1,000 h / 年試算)
- b. 着地精度やデータ品質の大幅な向上(現状、入力時の誤りのみ)

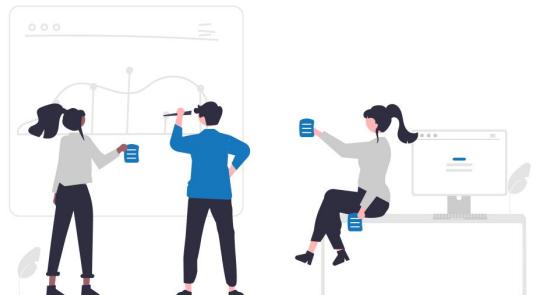
## 2. リカーリング プロセスの課題発見

## 3. 主要 KPI の変動要因分析 ✓

## 4. 仮説立案と行動後の迅速な検証 🏃

- a. 従来より数日-時に1週間以上、素早く行動検証が可能に

## 5. 成功モデルの横展開

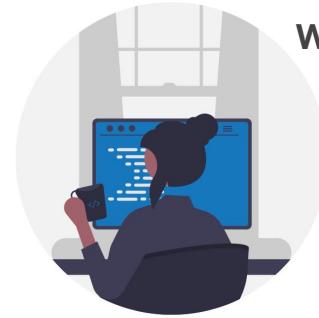


# 今後、データ活用組織として自走するために



## 組織学習 × 実践サイクル

primeNumber 社によるトレーニング  
Tableau の操作/活用の学習  
trocco® によるデータ処理の自動化



## 推進組織の専任体制化

データスチュワード

データエンジニア

Salesforce 開発エンジニア

詳細: <https://recruit.circu.co.jp/career/>

# 再掲：データ駆動経営に欠かせない3つの構成要素

## 1. データ エンジニアリング

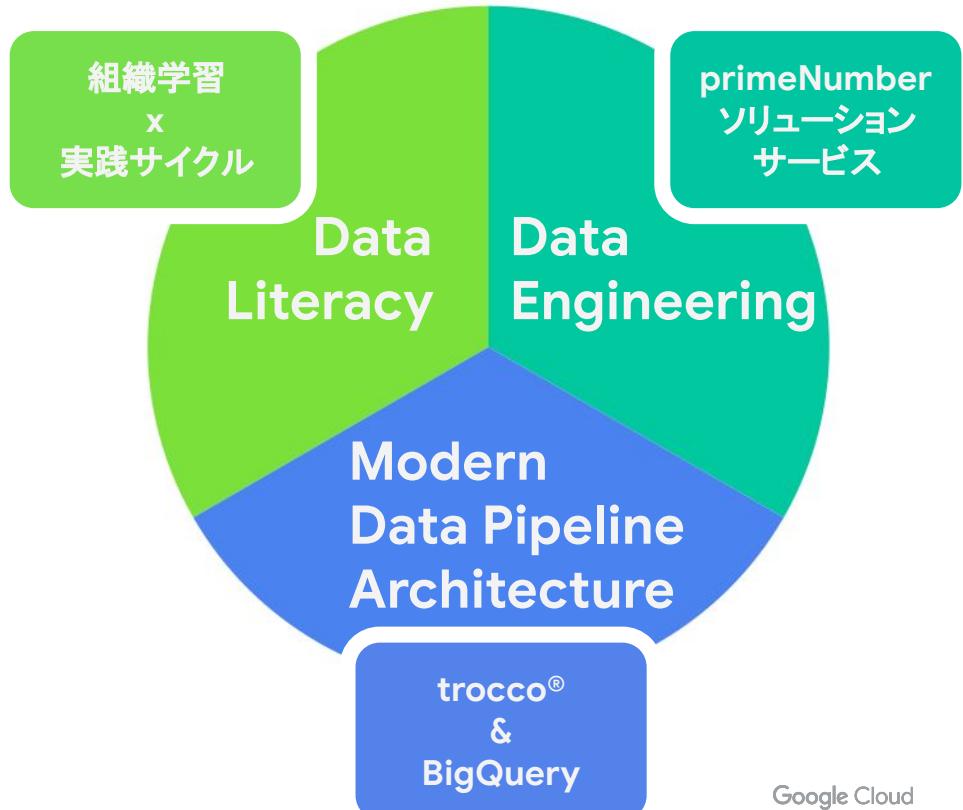
時には業務システムのデータの持ち方まで立ち返り  
ビジネスロジックや構造化データの見直しなど  
深いビジネス理解と多様な実現手腕を磨く

## 2. モダンなデータ分析パイプライン

省力開発・運用が可能なETL / DWH / BIツールの選定が  
日々の高速な仮説検証を支える

## 3. データリテラシーの組織学習

経営者をはじめ意思決定層に対する学習と実践のサイクル  
意思決定や業務の自動化・効率化を導く専任チームの組成



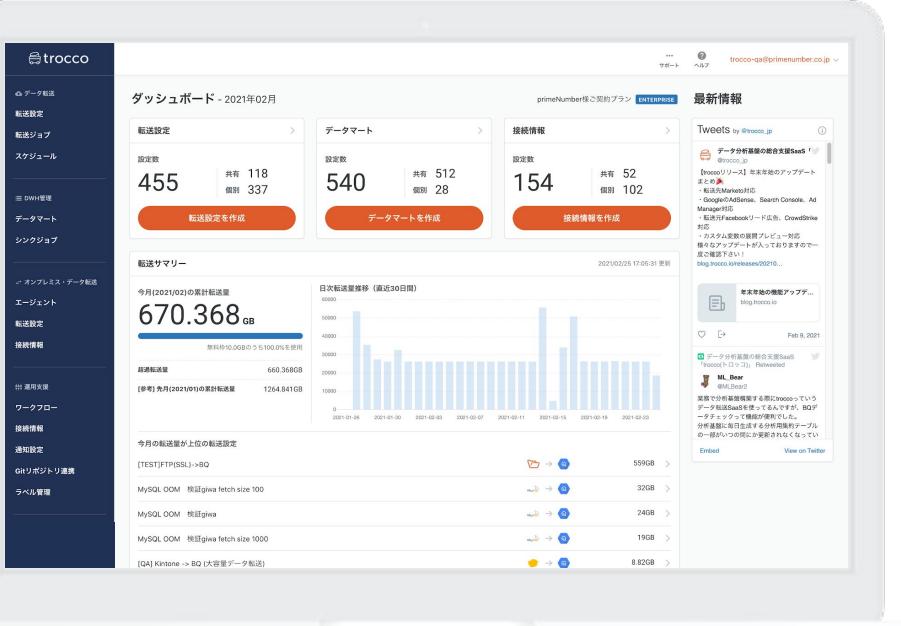
# trocco® って何？

各企業様のデータドリブンなアプローチを  
サポートする trocco® についてご説明します。



# trocco

# 国内最大級の導入実績を有する 日本発の「マネージド ETL & ワークフロー サービス」です



The screenshot shows the trocco dashboard interface. On the left, a sidebar lists various management options like Data Delivery, Delivery Job, Schedule, etc. The main area features three primary sections: 1) Delivery Status (455 total jobs, 118 pending, 337 completed), 2) Data Mart (540 total jobs, 512 pending, 28 completed), and 3) Integration Information (154 total jobs, 52 pending, 102 completed). Below these are two charts: a bar chart for daily delivery volume (Feb 1 to Feb 25) and a line chart for monthly cumulative delivery volume (Jan 2021 to Feb 2021). A detailed table at the bottom lists specific delivery configurations, such as MySQL OOM to BigQuery with a fetch size of 1000.



約 100 種のデータソースに対応した ETL



BigQuery 他、主要 DWH に対応した ELT



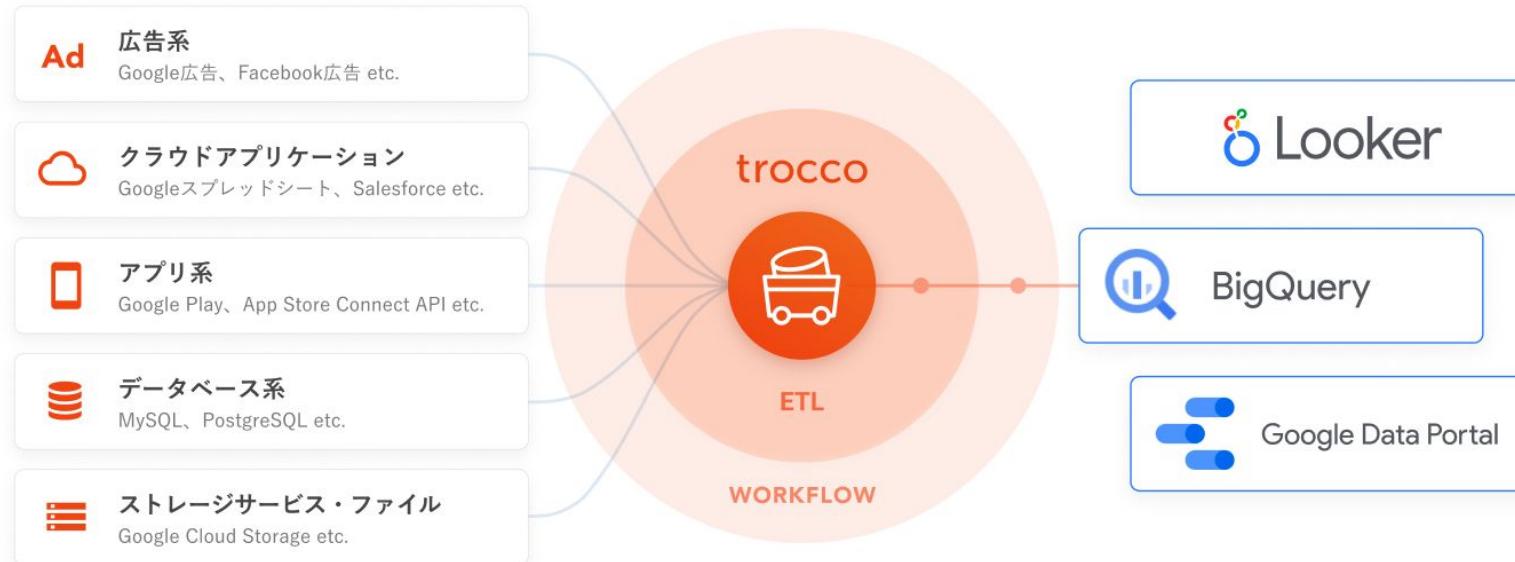
データ分析業務に適した ワークフロー

trocco® のミッション

データ エンジニアリングの泥臭い作業をなくし  
「攻めの仕事に  
集中できるように支援する」

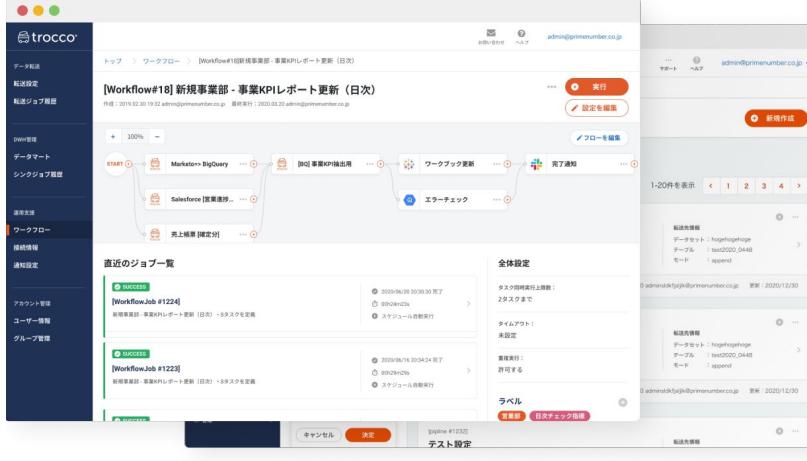


# 各種データソースから BigQuery の ETL / ELT 处理は trocco® で



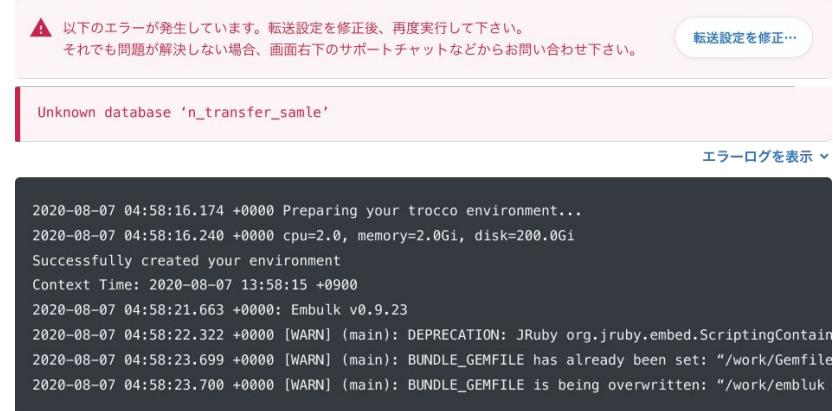
# 洗練されたUIで、 データ整備や分析環境を「素早く簡単に」実現できます

複雑なETLパイプラインも、  
わかりやすいGUIで定義



The screenshot displays the trocco interface for defining ETL workflows. On the left, a sidebar lists various management features like Data Mart, Sink Job Monitoring, Workflows, and User Management. The main area shows a detailed workflow diagram titled 'Workflow#18 新規事業部 - 事業KPIレポート更新 (日次)'. The diagram consists of several nodes: 'START', 'Marketo BigQuery', 'BigQuery [BQ] 事業KPI抽出用', 'ワークフック更新', 'エラーチェック', and 'Salesforce 事業連携'. Arrows indicate the flow between these nodes. A modal window is open in the center-right, titled '全体設定' (General Settings), containing fields for 'タスク間並列上限数' (Number of parallel tasks per task) set to '2タスクまで' (Up to 2 tasks), 'タイムアウト' (Timeout) set to '未設定' (Not set), and '重複実行' (Duplicate execution) set to '許可する' (Allow). At the bottom of the modal are 'キャンセル' (Cancel) and '次へ' (Next) buttons.

エラー要約付きのリアルタイムログで  
迅速な対応が可能



The screenshot shows the trocco UI handling an error. At the top, a red warning icon and message say: '以下のエラーが発生しています。転送設定を修正後、再度実行して下さい。それでも問題が解決しない場合、画面右下のサポートチャットなどからお問い合わせ下さい。' (An error has occurred. Please correct the transfer settings and re-execute. If the problem persists, contact support via the chat at the bottom right.) Below this, a red box highlights the error message: 'Unknown database 'n\_transfer\_sample''.

On the right, a button says '転送設定を修正...' (Correct transfer settings...). At the bottom, a large black box displays the error log:

```

2020-08-07 04:58:16.174 +0000 Preparing your trocco environment...
2020-08-07 04:58:16.240 +0000 cpu=2.0, memory=2.0Gi, disk=200.0Gi
Successfully created your environment
Context Time: 2020-08-07 13:58:15 +0900
2020-08-07 04:58:21.663 +0000: Embulk v0.9.23
2020-08-07 04:58:22.322 +0000 [WARN] (main): DEPRECATION: JRuby org.jruby.embed.ScriptingContainer
2020-08-07 04:58:23.699 +0000 [WARN] (main): BUNDLE_GEMFILE has already been set: "/work/Gemfile"
2020-08-07 04:58:23.700 +0000 [WARN] (main): BUNDLE_GEMFILE is being overwritten: "/work/embulk

```

# データマートELT(BigQuery 格納後処理)も機能が豊富

ホーム > データマート定義一覧 > HCの対象Forecastデータのみを抽出 > データマート定義の編集

① 現在選択中です。確認画面で「適用」または「下書き保存」を行うままで、変更是保存されません。

### データマート定義の編集 - datamart#4320 HCの対象Forecastデータのみを抽出

 BigQuery

変数名 [必須] データ型 [必須]  
\$ym\$ 時刻・日付 (キューイング時)  
値 [必須] 単位 [必須] 前/後 [必須] 日付フォーマット [必須] タイムゾーン [必須]  
1 ケ月 前 %Y%m Asia/Tokyo (+09:00)

変数名 [必須] データ型 [必須]  
\$yyyymm\$ 時刻・日付 (キューイング時)  
値 [必須] 単位 [必須] 前/後 [必須] 日付フォーマット [必須] タイムゾーン [必須]  
1 ケ月 前 %V-%m Asia/Tokyo (+09:00)



データの抽出を行い  
新しいマートを構築

この項目に関するヘルプ

クエリ設定

③ クエリ実行モード

 データ転送モード  
SQLと出力先テーブルを指定するだけで、簡単にテーブルへの問い合わせ・追記を実行することができます。

 自由起述モード  
接続先のDWHに対して、任意のクエリ (DDL, DELETE, INSERTなど) を自由に実行することができます。

クエリ [必須] カスタム変数用引数

```

1 -- BQMLモデルの作成
2 CREATE OR REPLACE MODEL sfdc_demo.vgboost_model
3 OPTIONS(
4   MODEL_TYPE='boosted_tree_classifier', INPUT_LABEL_COLS=["buy_LT"]
5 ) AS
6 SELECT
7   EXCEPT(user_id, Name, Age, opp, company, todo)
8   FROM
9   `prd_m1.sfdc_demo.train_age`
10 ;
11
12 -- BI用データスナップショットを作成する
13 DROP SNAPSHOT TABLE IF EXISTS `prd_snap.bi_view_fromSFDC_$date$`;
14 CREATE SNAPSHOT TABLE `prd_snap.bi_view_fromSFDC_$date$`
15 CLONE `prd_snap.bi_view_fromSFDC_$date$`;

```

④ プレビュー実行

クエリ [必須] カスタム変数用引数

```

1 (
2   SELECT
3     Create_Date,
4     Target_Date,
5     Team,
6     CAST(First_Day_Month_Budget_Conslt_HC AS int) AS First_Day_Month_Budget_Conslt_HC,
7     CAST(First_Day_Month_Non_Budget_Conslt_HC AS int) AS First_Day_Month_Non_Budget_Conslt_HC,

```

ホーム > データマート定義一覧 > 学習モデルを作成 > データマート定義の編集

② 現在選択中です。確認画面で「適用」または「下書き保存」を行うままで、変更是保存されません。

### データマート定義の編集 - datamart#3399 学習モデルを作成

 trocco

データの操作が可能  
BQ が許容するクエリにも対応

この項目に関するヘルプ

データ転送モード

SQL出力先テーブルを指定するだけで、簡単にテーブルへの問い合わせ・追記を実行することができます。

自由起述モード

接続先のDWHに対して、任意のクエリ (DDL, DELETE, INSERTなど) を自由に実行することができます。

クエリ [必須] カスタム変数用引数

```

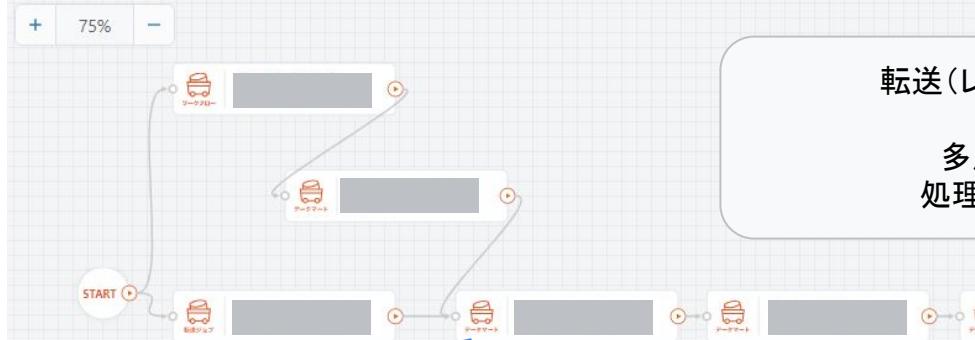
1 -- BQMLモデルの作成
2 CREATE OR REPLACE MODEL sfdc_demo.vgboost_model
3 OPTIONS(
4   MODEL_TYPE='boosted_tree_classifier', INPUT_LABEL_COLS=["buy_LT"]
5 ) AS
6 SELECT
7   EXCEPT(user_id, Name, Age, opp, company, todo)
8   FROM
9   `prd_m1.sfdc_demo.train_age`
10 ;
11
12 -- BI用データスナップショットを作成する
13 DROP SNAPSHOT TABLE IF EXISTS `prd_snap.bi_view_fromSFDC_$date$`;
14 CREATE SNAPSHOT TABLE `prd_snap.bi_view_fromSFDC_$date$`
15 CLONE `prd_snap.bi_view_fromSFDC_$date$`;

```

カスタム変数展開結果を確認

④ ジョブ起動設定

## サーキュレーション様のワークフローイメージ



The screenshot shows a workflow diagram with various nodes representing different actions or data sources. A blue arrow points from the workflow area to a Google Cloud Platform interface.

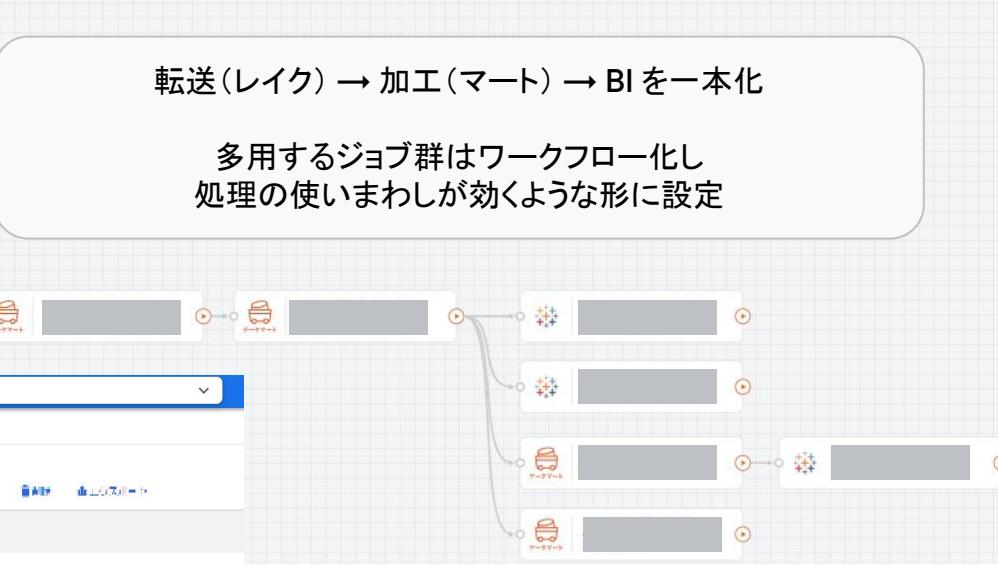
**Google Cloud Platform Interface:**

- Search bar: 検索 プロダクト、リソース、ドキュメント (/)
- Navigation: 機能と情報, ショートカット, エディタタブを無効にする
- Project Selector: 上スクエア - 1K
- Data Studio: **side\_user** (eri, 2023/08/24)
- Table View: side\_user (データを表示する) - 1行

名前	詳細	操作		
pl_Pardon_UserId_c	Managed	Id	UserType	Username
11	0057E00000Qp1404K	0057E00000WtVAV	Standard	jiro.yamada@concur.jp
12	0057E00000Qp1404K	0057E00000WtVAV	Standard	isamu.watanabe@concur.jp
13	0057E00000Qp1404K	0057E00000WtVAV	Standard	isamu.haneda@concur.jp
14	0057E00000Qp1404K	0057E00000Qp1404K	Standard	jiro.yamada@concur.jp
15	0057E00000Qp1404K	0057E00000WtVAV	Standard	isamu.watanabe@concur.jp
16	0057E00000Qp1404K	0057E00000WtVAV	Standard	jiro.yamada@concur.jp
17	XXXXXXXXXX	-	Standard	jiro.yamada@concur.jp
18	0057E00000Qp1404K	0057E00000WtVAV	Standard	isamu.watanabe@concur.jp
19	0057E00000Qp1404K	0057E00000WtVAV	Standard	jiro.yamada@concur.jp

**転送(レイク) → 加工(マート) → BI を一本化**

多用するジョブ群はワークフロー化し  
処理の使いまわしが効くような形に設定



BigQuery 側に適切なデータ断面を残すため  
幕等性を持った設計やスナップショットを取得

# trocco® が事業に寄与できること

## 営業管理(本日のお話はこちら)

- 営業活動評価のための複数ツールデータ統合
- 営業数値の可視化
- MA × SFA の連携

## マーケティング領域

- チャネルの貢献度評価(LTV、ROAS 等)
- マーケティング分析時間の創出
- マーケティング データの統合

## 事業分析・グロース

- UI / UX 改善につながるユーザ利用分析
- 事業トレンドを把握するスナップショット分析

## 業務効率化

- 提供サービスの一部業務自動化
- エンジニアリソースの最適配置
- データ分析環境の運用標準化・安定化

# 月額基本料金 + 使った分だけ従量課金でご利用いただけます

## Light プラン

(月間)総転送時間の目安

250 時間

※コネクタ数制限あり

基本料金

月額 10 万円～

## Standard プラン

(月間)総転送時間の目安

1250 時間

※コネクタ数制限なし

基本料金

月額 30 万円～

## Enterprise プラン

(月間)総転送時間の目安

ご相談ベース

※コネクタ数制限なし

基本料金

随時お見積もり

## オプション機能

- ✓ グループ機能
- ✓ 多言語プログラミング ETL
- ✓ Git 設定変更管理
- ✓ エージェント機能
- ✓ API 機能
- ✓ Web 行動ログ収集 SDK

(2022年4月時点)最新情報は製品ページをご確認ください。<https://trocco.io/lp/plan.html>



trocco®についてもっと知りたい方は、

”trocco ETL”で検索 or QR 読み取りで 製品紹介ページへ



trocco ETL



BigQuery と trocco® の組み合わせのご提案など  
貴社に最適なデータ分析基盤構築のご相談も承ります。



# Thank you.

