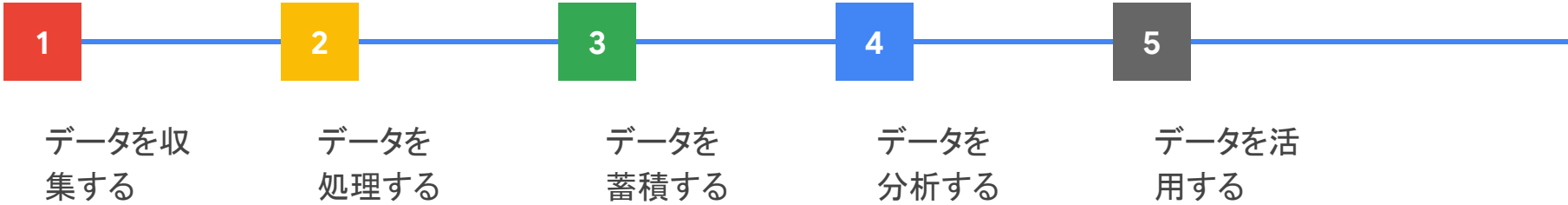
10 分で分かる
Google Cloud データ分析関連サービス

Google Cloud
カスタマーエンジニア
星 美鈴

Google Cloud

データ活用の流れ



Google Cloud から提供されるデータ分析関連プロダクト

収集

スケーラブルな
データ収集

- Cloud Pub/Sub
- Data Transfer Service
- Cloud IoT Core
- Storage Transfer Service

処理

信頼性のある
ストリーム データ
パイプライン

- Cloud Dataflow
- Cloud Dataproc
- Cloud Dataprep

蓄積

データレイク
データ ウェアハウス

- Cloud Storage
- **BigQuery** storage

分析

データ ウエアハウス

- **BigQuery** analysis engine

活用

先進的な分析

- Cloud AI Services
- Google Data Portal
- Looker
- Connected Sheets

全体

- Cloud Composer
- Cloud Data Fusion
- Data Catalog

Google Cloud から提供されるデータ分析関連プロダクト

収集

スケーラブルな
データ収集

- Cloud Pub/Sub
- Data Transfer Service
- Cloud IoT Core
- Storage Transfer Service

処理

信頼性のある
ストリーム データ
パイプライン

- Cloud Dataflow
- Cloud Dataproc
- Cloud Dataprep

蓄積

データレイク
データ ウェアハウス

- Cloud Storage
- **BigQuery** storage

分析

データ ウエアハウス

- **BigQuery** analysis engine

活用

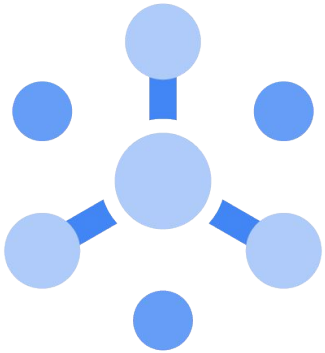
先進的な分析

- Cloud AI Services
- Google Data Portal
- Looker
- Connected Sheets

全体

- Cloud Composer
- Cloud Data Fusion
- Data Catalog

Cloud Pub/Sub

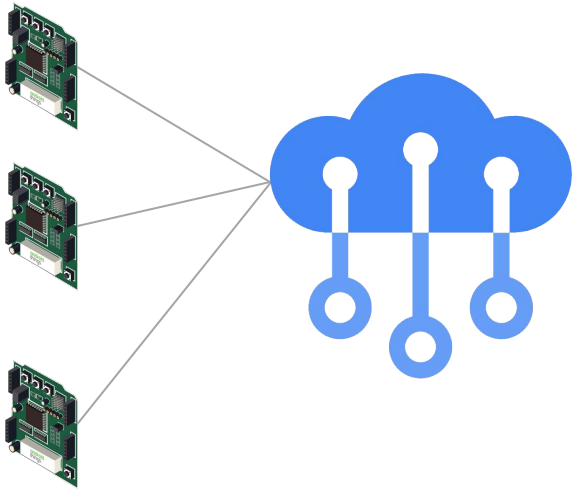


システム間の疎結合化/長時間接続を提供する
マネージドメッセージングサービス

0~数百 GB/sに自動スケール

多様な配信パターンに対応
（一対一 / 一対多 / 多対多、Push / Pull）

Cloud IoT Core



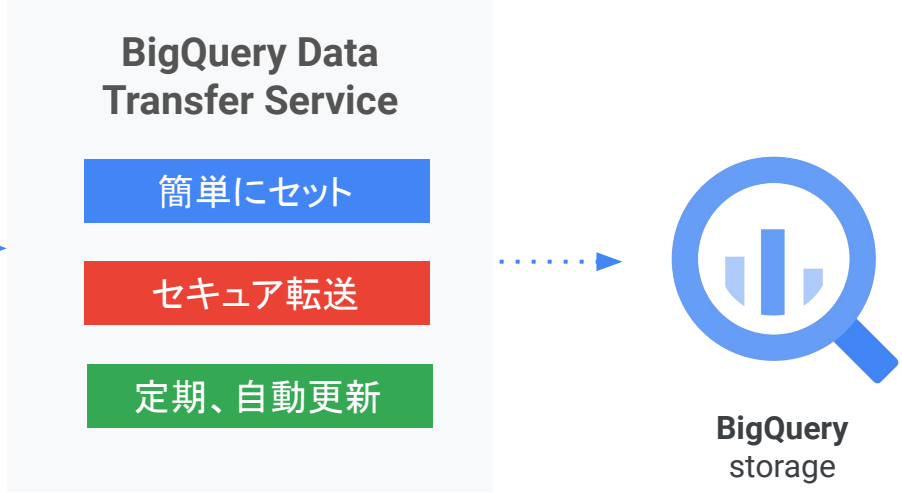
多数のデバイス管理が可能な
マネージドサービス

セキュアに双方向の通信が可能

BigQuery Data Transfer Service

サポートされるデータソース

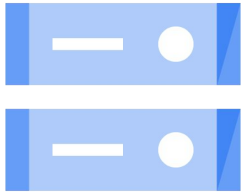
- Amazon S3
- キャンペーンマネージャ
- Cloud Storage
- Google 広告
- Googleアドマネージャー
- Google Merchant Center
- Google Play
- 検索広告 360
- YouTube チャンネル
- YouTube コンテンツ所有者
- Teradata
- Amazon Redshift
- その他サードパーティ



Storage Transfer Service

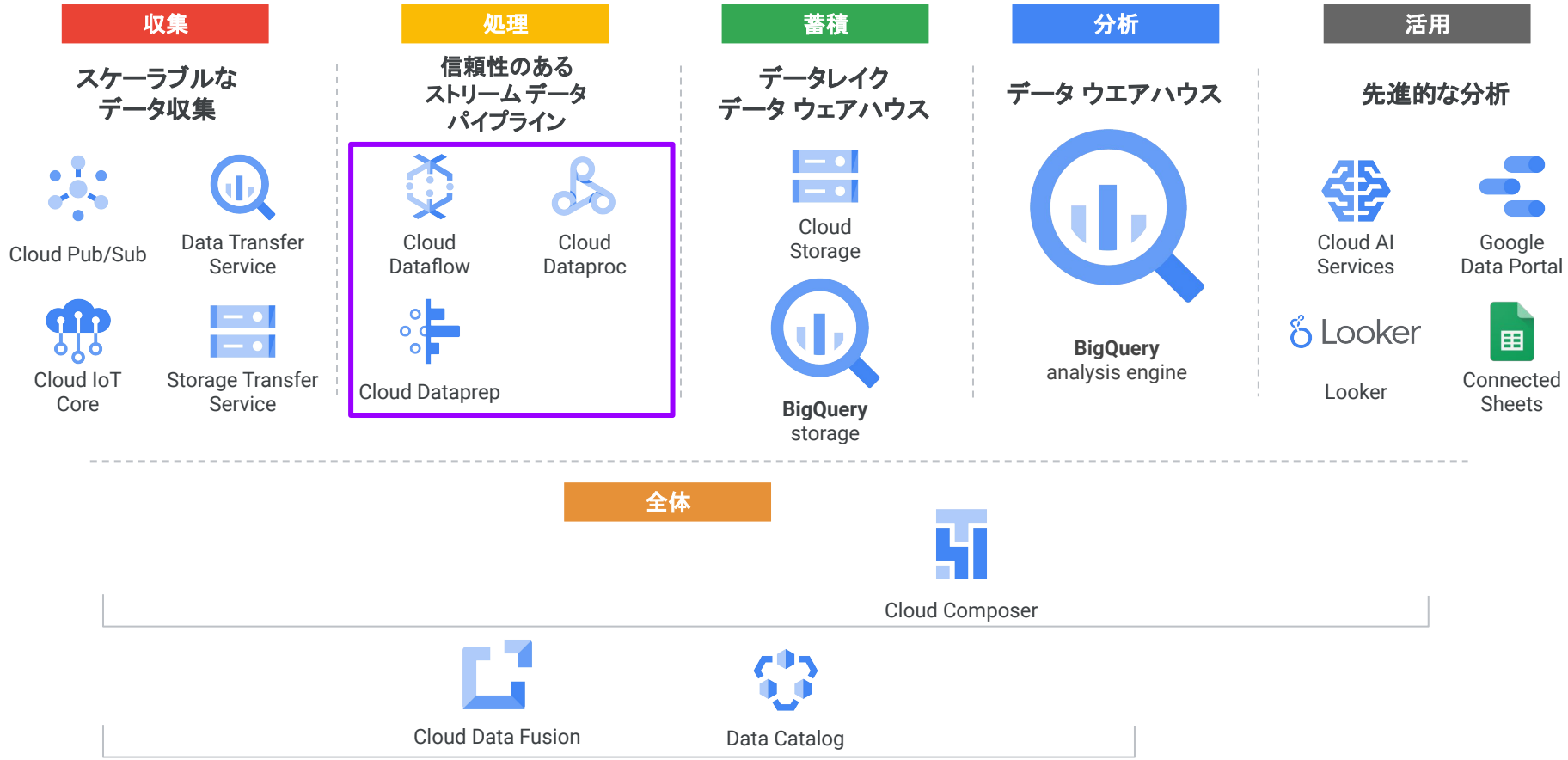
サポートされるデータソース

- Amazon S3
- Azure Blob Storage
- バケットの異なるCloud Storage
- ウェブアドレスで参照できるデータ
- オンプレミスデータ

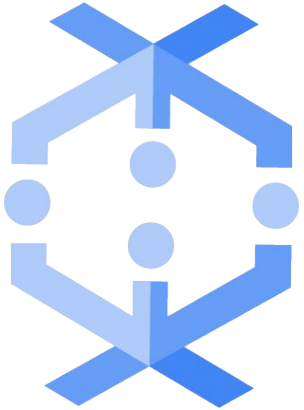


Cloud Storage

Google Cloud から提供されるデータ分析関連プロダクト

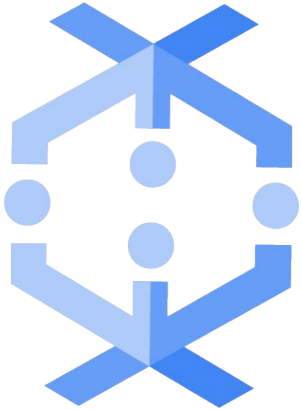


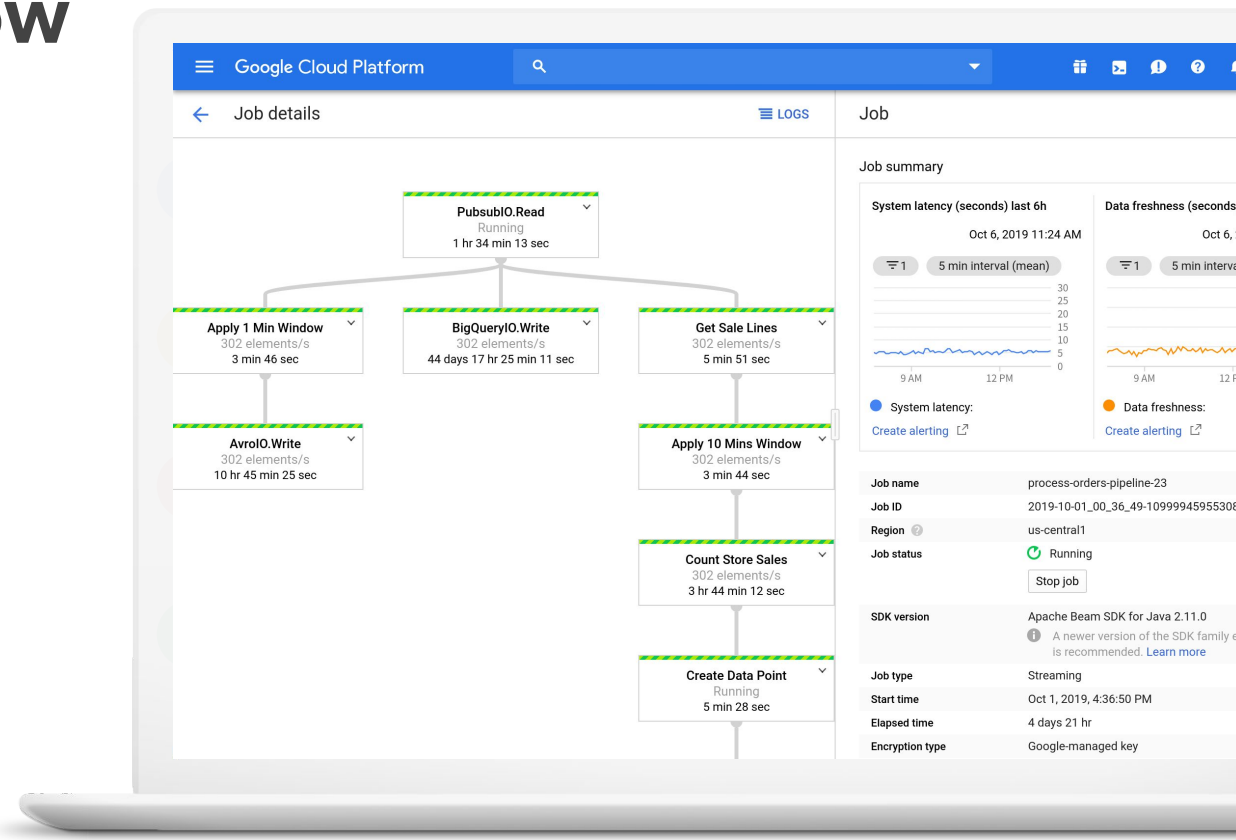
Cloud Dataflow



- フルマネージドのデータ処理サービス
- オープン（Apache Beam）
- 分散処理、オートスケール
- ストリームおよびバッチの両方に対応

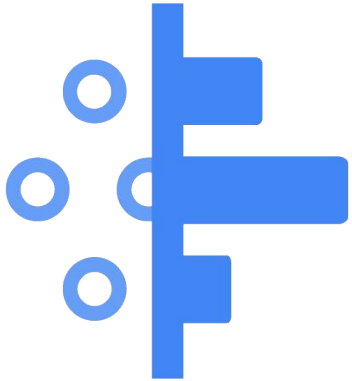
Cloud Dataflow



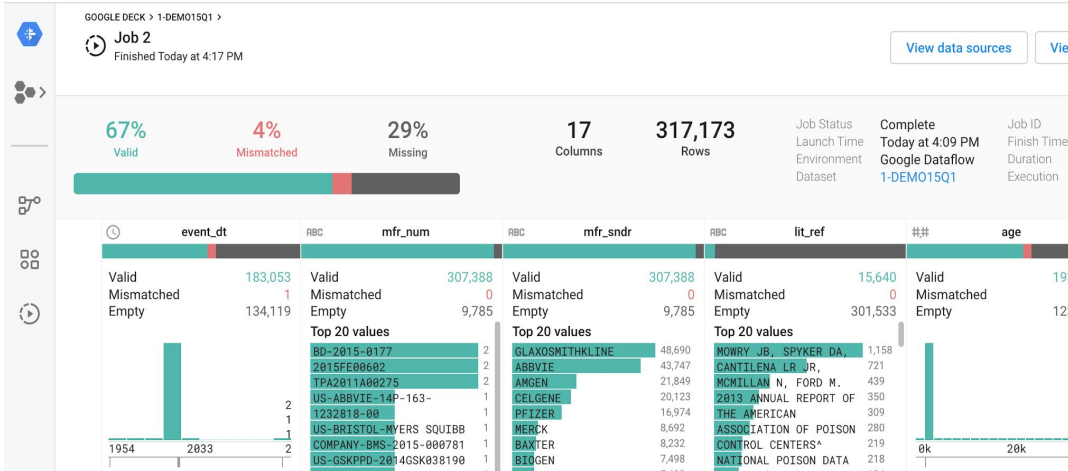


Google Cloud

Cloud Dataprep

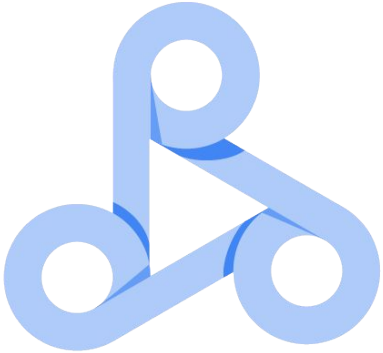


- フルマネージドのデータ処理サービス
- GUI でデータを整形、加工可能



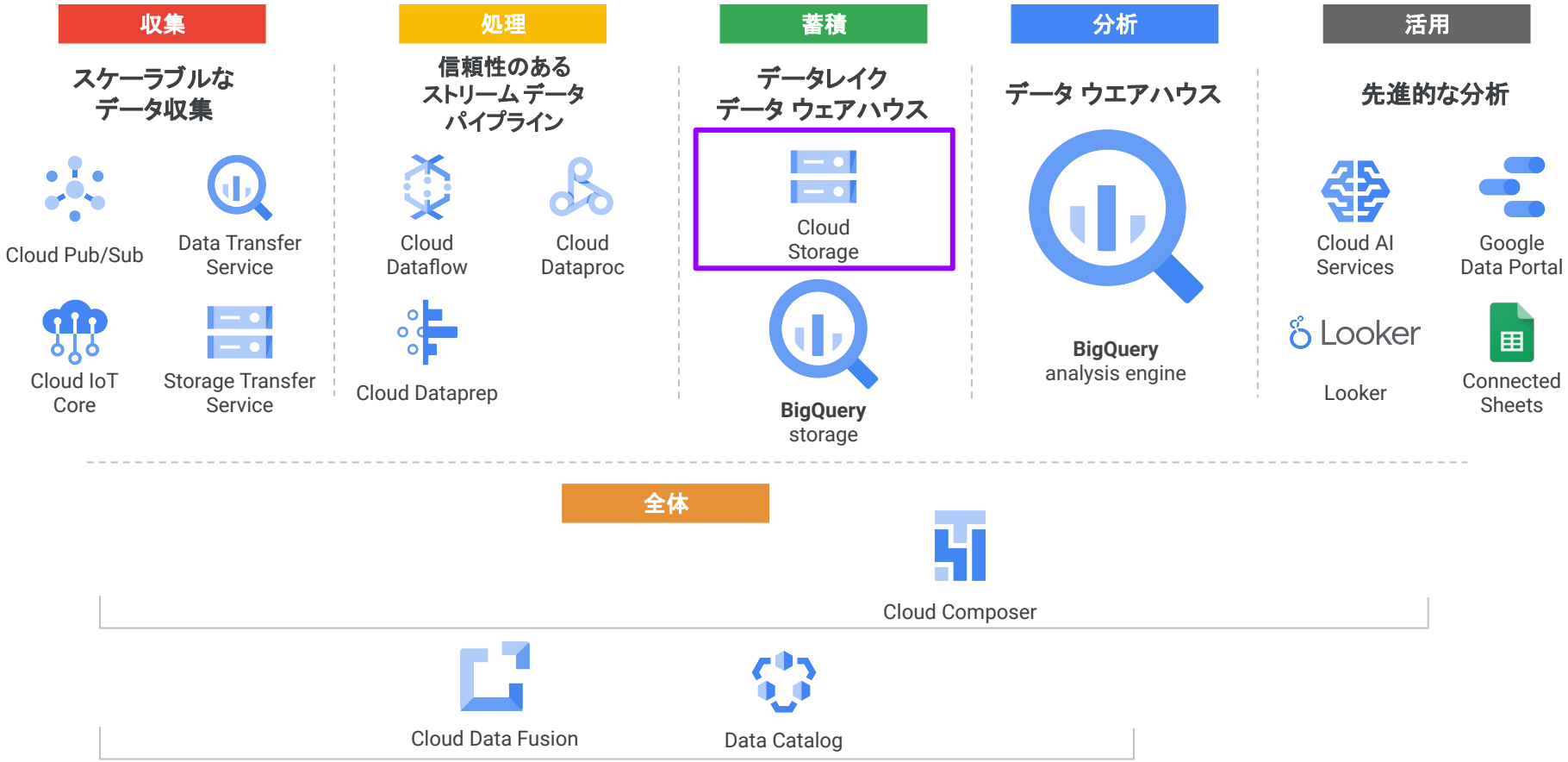
Google Cloud

Cloud Dataproc

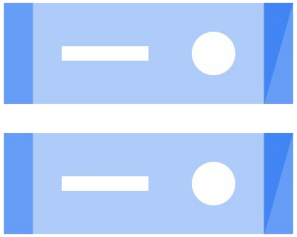


- Apache Hadoop / Spark のマネージドサービス
- ワンクリックまたはワンコマンドにより約 90 秒で Hadoop クラスタを構築
- プリエンプティブル VM の利用や不要時の無効化によりコストを最適化

Google Cloud から提供されるデータ分析関連プロダクト

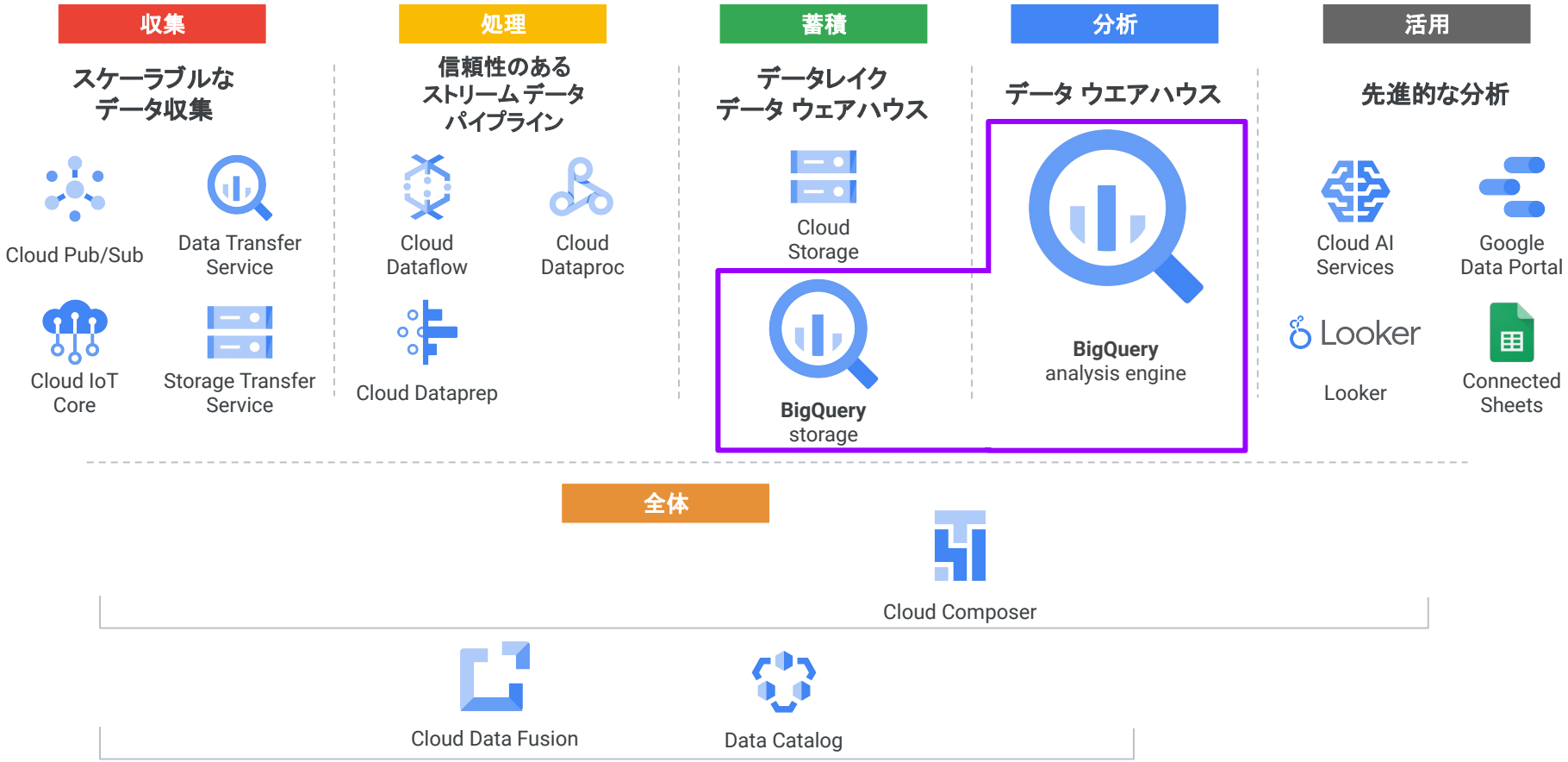


Cloud Storage



- 容量無制限のオブジェクトストレージ
- 高い耐久性（年間 99.999999999 %）
- アクセス頻度に応じたストレージクラス（Standard / Nearline / Coldline / Archive）
- 指定条件に応じ低料金クラスへ容易に移行

Google Cloud から提供されるデータ分析関連プロダクト



Google BigQuery

エンタープライズ向け
データウェアハウス

無制限のストレージと
ペタバイト級クエリ

強固なセキュリティと
高可用性

特徴

フルマネージド、サーバーレス

リアルタイム分析

AI の基盤 / BigQuery ML

BI の基盤 / BI Engine

Google Cloud から提供されるデータ分析関連プロダクト

収集

スケーラブルな
データ収集

- Cloud Pub/Sub
- Data Transfer Service
- Cloud IoT Core
- Storage Transfer Service

処理

信頼性のある
ストリーム データ
パイプライン

- Cloud Dataflow
- Cloud Dataproc
- Cloud Dataprep

蓄積

データレイク
データ ウェアハウス

- Cloud Storage
- **BigQuery** storage

分析

データ ウエアハウス

- **BigQuery** analysis engine

活用

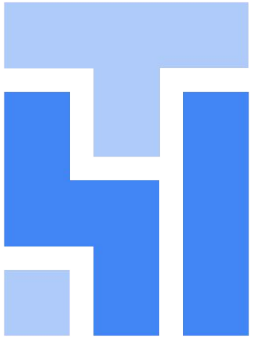
先進的な分析

- Cloud AI Services
- Google Data Portal
- Looker
- Connected Sheets

全体

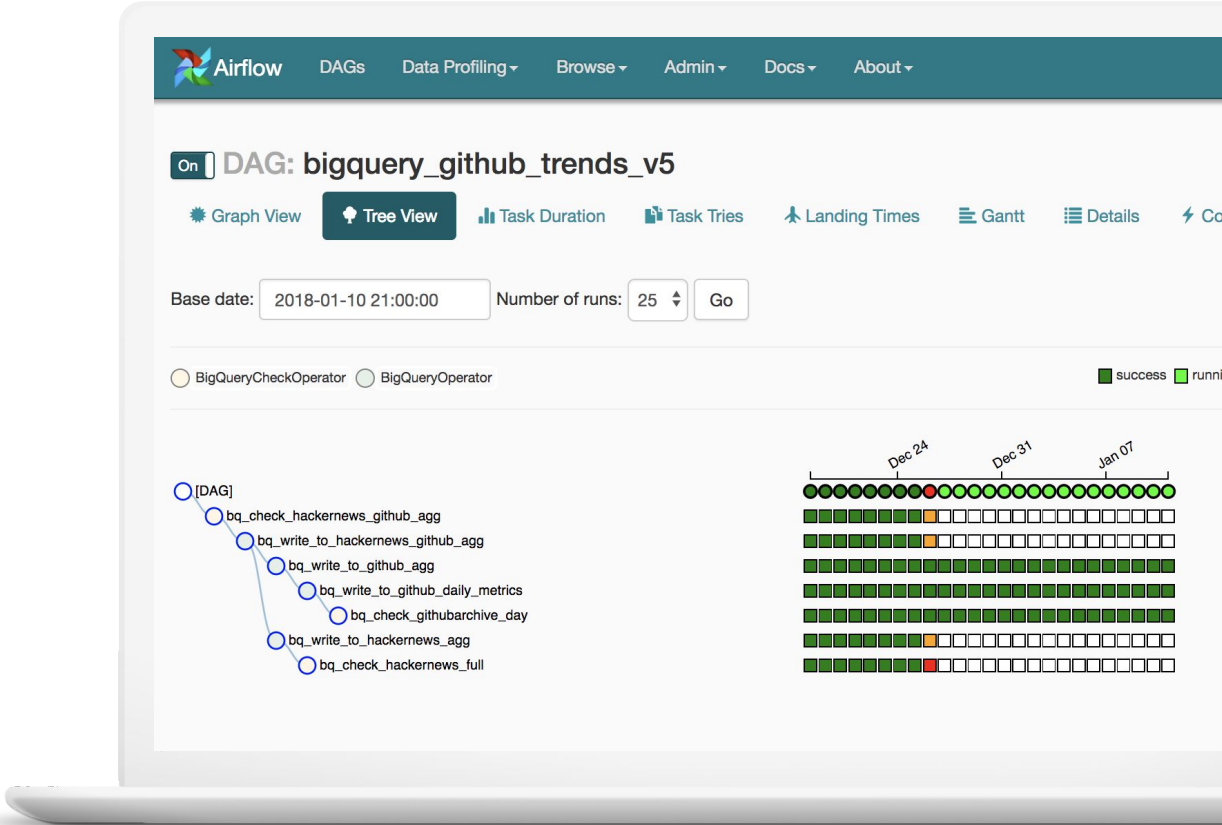
- Cloud Composer
- Cloud Data Fusion
- Data Catalog

Cloud Composer



- マネージドのデータ系ジョブ管理サービス
- オープン（Apache Airflow）
- ハイブリッド、マルチクラウドに跨がるパイプラインを作成可能

Cloud Composer

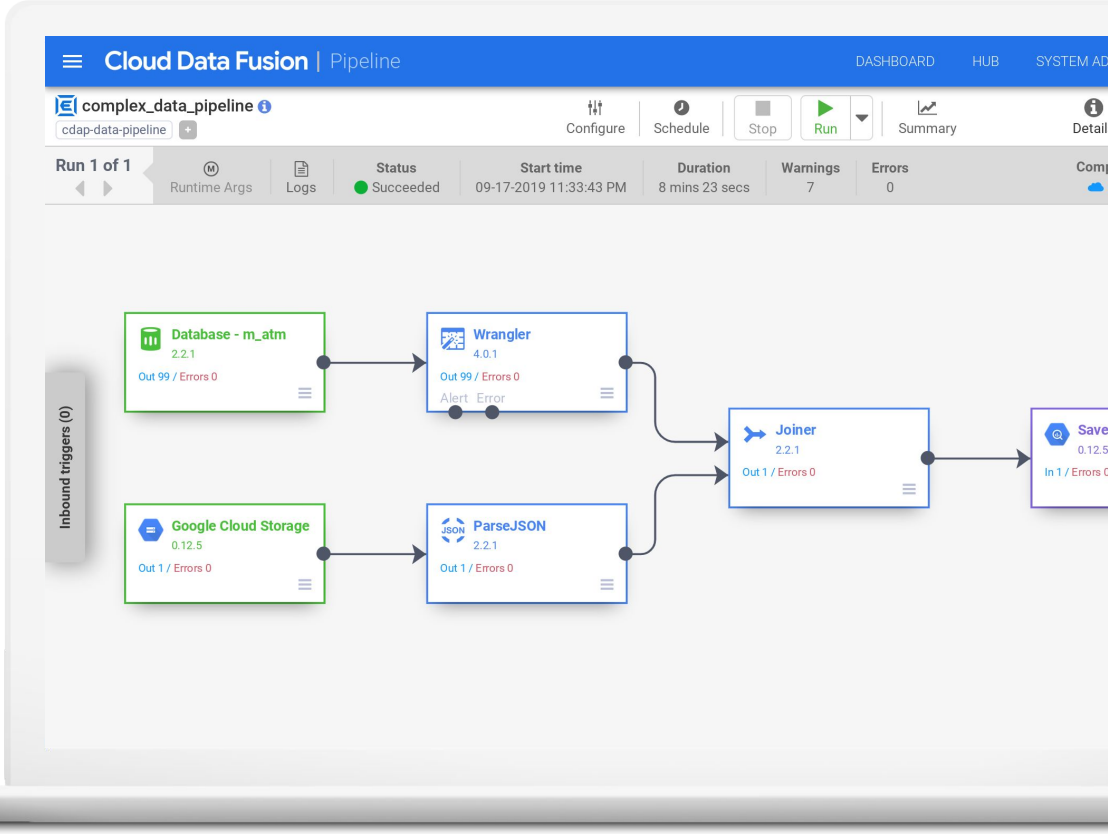


Google Cloud

Cloud Data Fusion

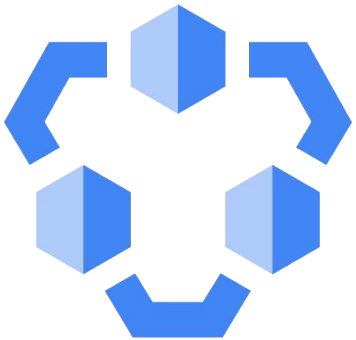
- マネージドのデータ統合管理サービス
- オープン（CDAP）
- GUI 操作でデータ処理、メタデータ管理、ハイブリッド/マルチクラウドに跨がるパイプラインの管理が可能

Cloud Data Fusion



Google Cloud

Data Catalog



- マネージドのメタデータ管理サービス
- 検索技術によるデータ資産検出の効率化
- タグ付けなどによるデータ管理の簡易化

Google Cloud から提供されるデータ分析関連プロダクト

収集

スケーラブルな
データ収集

- Cloud Pub/Sub
- Data Transfer Service
- Cloud IoT Core
- Storage Transfer Service

処理

信頼性のある
ストリーム データ
パイプライン

- Cloud Dataflow
- Cloud Dataproc
- Cloud Dataprep

蓄積

データレイク
データ ウェアハウス

- Cloud Storage
- **BigQuery** storage

分析

データ ウエアハウス

- **BigQuery** analysis engine

活用

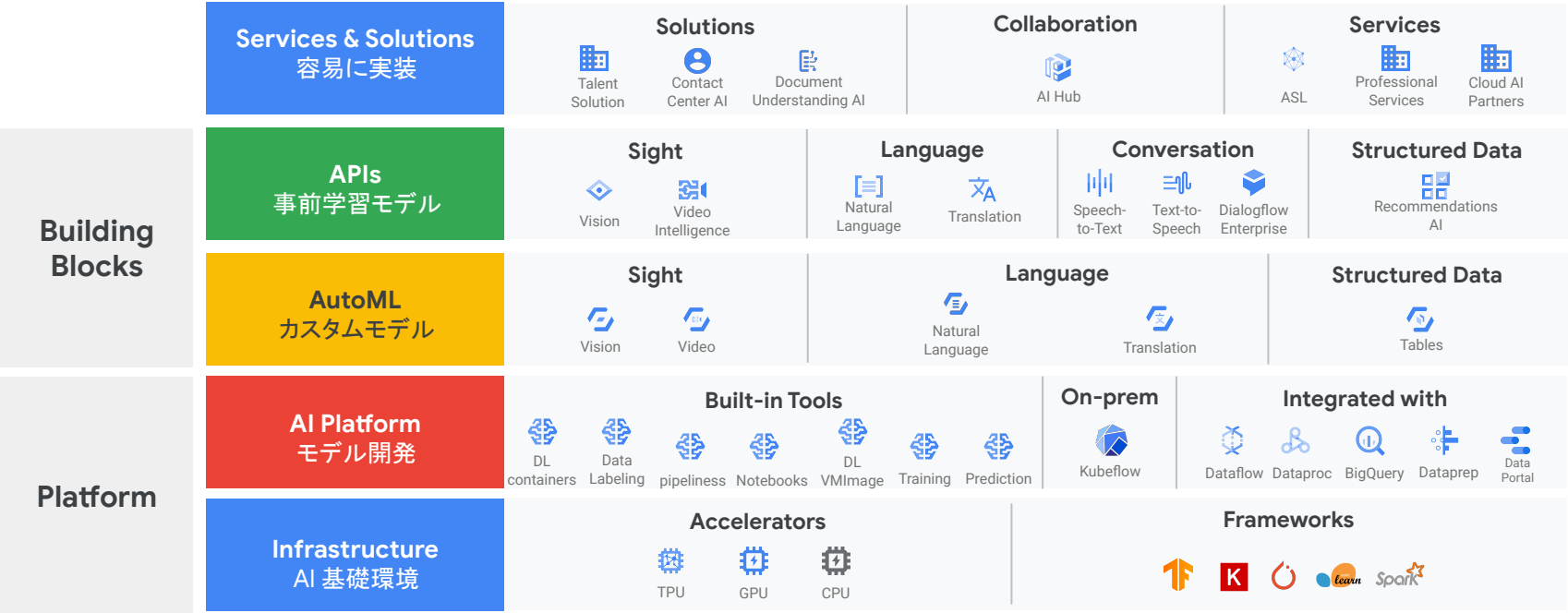
先進的な分析

- Cloud AI Services
- Google Data Portal
- Looker
- Connected Sheets

全体

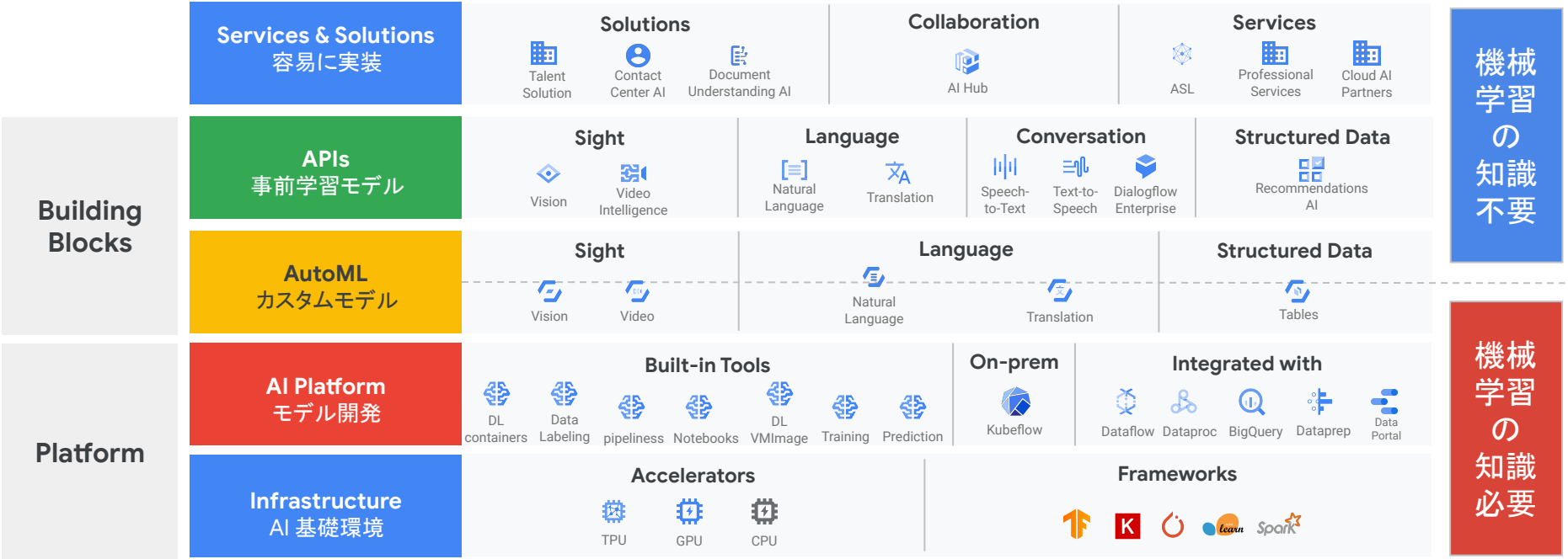
- Cloud Composer
- Cloud Data Fusion
- Data Catalog

Cloud AI サービス群



Google Cloud

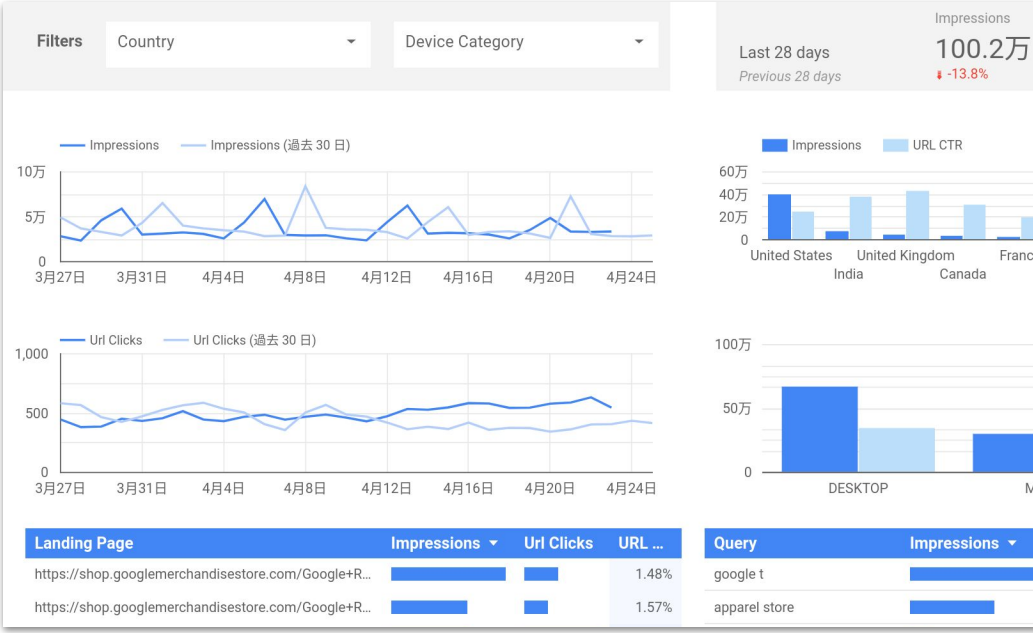
Cloud AI サービス群



Google Cloud

Data Portal

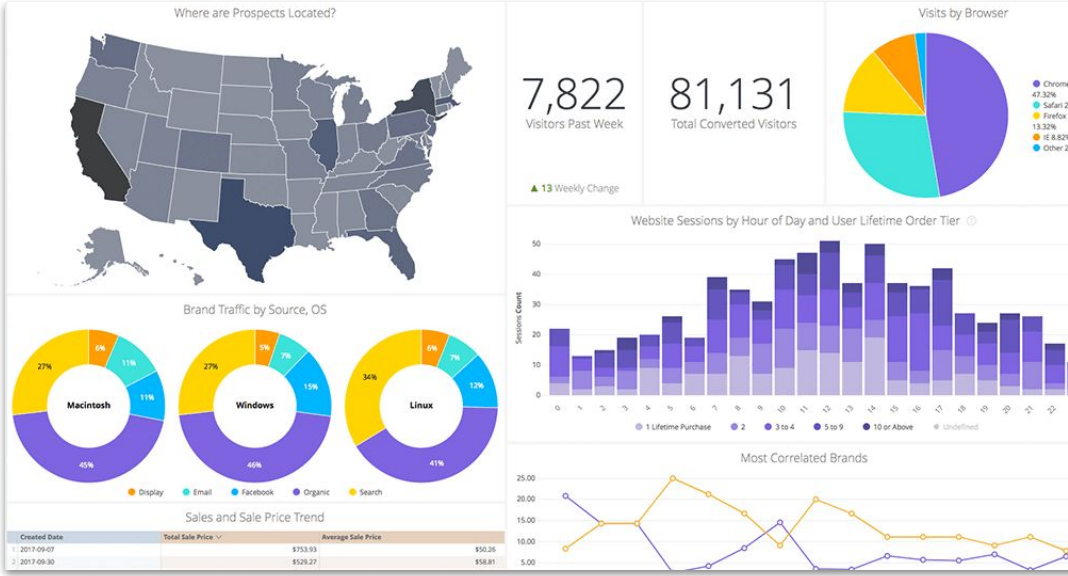
- 無料のデータ可視化ツール
- アドホックな可視化や
  レポートに最適
- 作成レポートを他ユーザ、
  グループと共有可能



Google Cloud

Looker

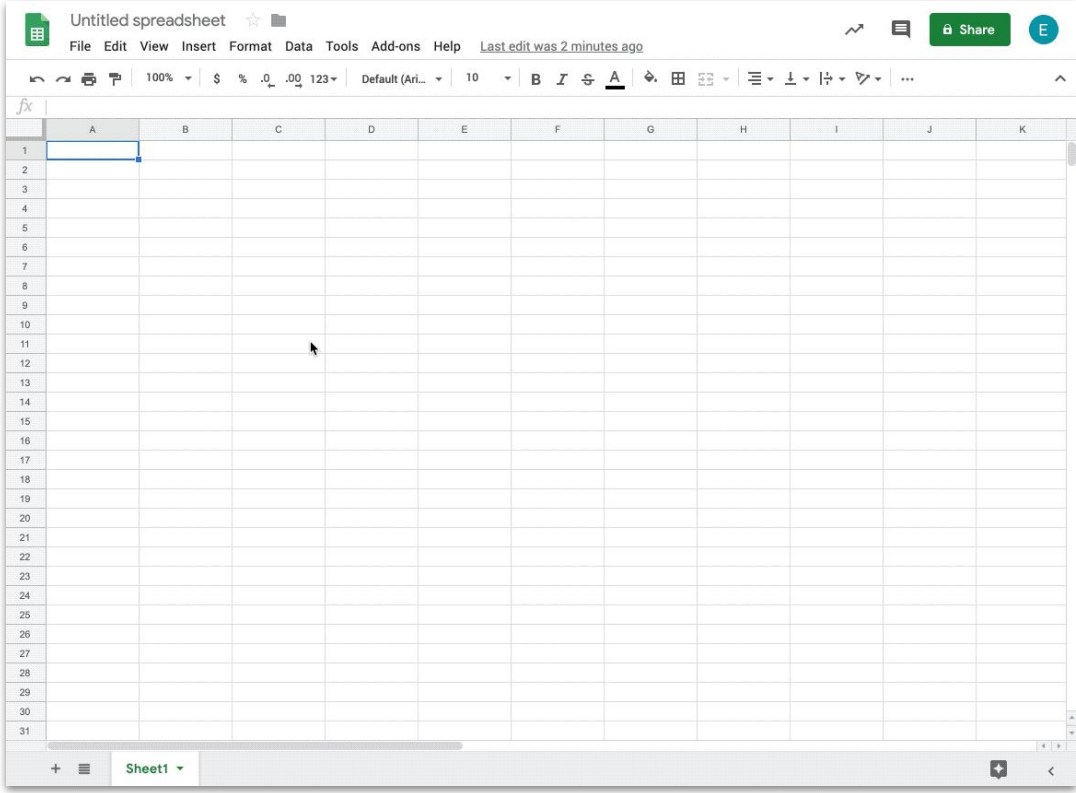
- エンタープライズ向け
  BI 分析プラットフォーム
- セキュアで、企業全体の
  ガバナンスを維持可能
- ビジネスツールと連携し
  インサイトを迅速に行動へ



Google Cloud

Connected Sheets

- BigQuery の何十億行のデータを Sheets で分析
- データは常に最新を維持
- 慣れ親しんだスプレッドシートでデータ探索が可能



※ Google Workspace Enterprise の契約が必要

Google Cloud

Google Cloud データ分析関連プロダクトをひと言で表現

収集	Cloud Pub/Sub	リアルタイムにデータを処理するマネージドメッセージキュー
	Cloud IoT Core	IoT デバイスの管理やデータ取り込みのためのマネージドサービス
	Data Transfer Service	BigQuery にデータを取り込むサービス
	Storage Transfer Service	Google Cloud Storage にデータを取り込むサービス
処理	Cloud Dataflow	ストリーム / バッチでの ETL 処理が可能なマネージドApache Beam
	Cloud Dataprep	GUI で ETL 処理が可能なマネージドサービス
	Cloud Dataproc	マネージドSpark / Hadoop
蓄積	Cloud Storage	オブジェクトストレージ
分析	BigQuery	ビッグデータ用分析システム、超並列分散データウェアハウス
全体	Cloud Composer	ワークフローオーケストレーションのためのマネージドApache Airflow
	Cloud Data Fusion	GUI でデータパイプラインを管理するマネージドCDAP
	Data Catalog	メタデータ管理のためのマネージドサービス
活用	Cloud AI	ユーザの専門知識レベルに幅広く対応したAI サービス群
	Google Data Portal	データを簡単に可視化するダッシュボード
	Looker	エンタープライズ向けBI プラットフォーム
	Connected Sheets	Google スプレッドシートからBigQuery へアクセスし分析する機能

データ基盤全体をカバーする
書籍が 2/20 に出版されました！

是非皆様のデータ活用に
お役立てください！

Google Cloud

Thank you

Google Cloud