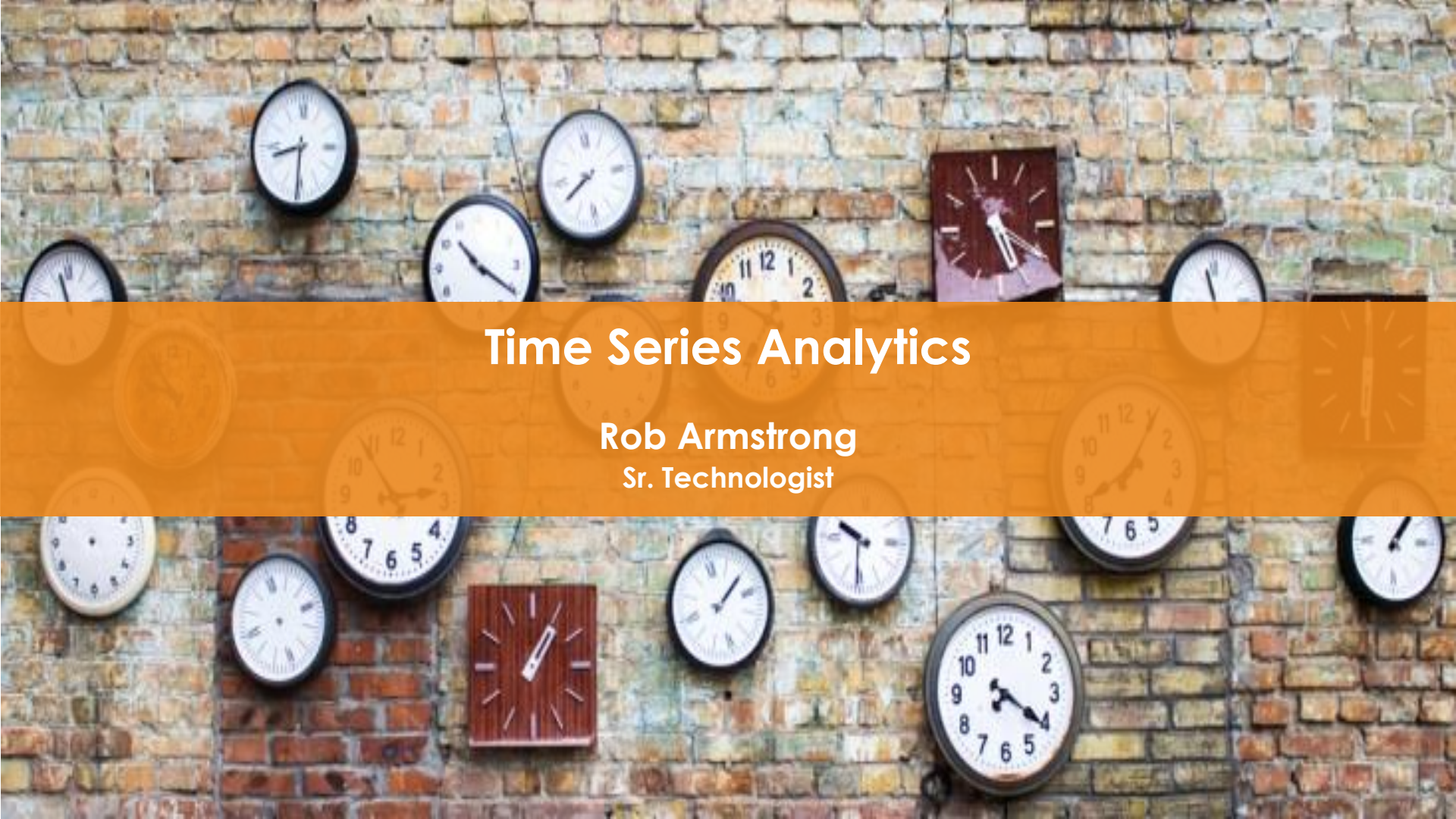


Time Series Analytics

Rob Armstrong
Sr. Technologist

Key Takeaways

- Integrate new data sources and include new user communities for broader insights
- Simplified and optimized capabilities to accelerate time based analytics
- Leverage integrated analytics with the scale, optimization, and management of Teradata

Business enabled by Time Series Functions

Business users, Operational Technology engineers & data scientists

Joining sensor data to corporate contextual data

Scalability across manufacturing plants, oil fields, fleets of vehicles, power grids, …

Solving complex analysis previously unsolvable

Optimized & parallel performance

Enabling IoT Analytics

The operationalization of real time Operational Technology data from sensors/IoT/RFID and other sources:

Deep Sea Oil & Gas Production…

Operational Failures:

Life Threatening Conditions &

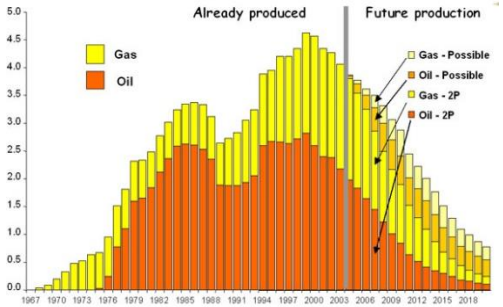
Lost Revenues

Preventing equipment failure:

- Know Safety Tolerances are Reached
- Safer Work Environment

Predicting Changes to Sub-Surface Conditions with Time Series Data:

- Ensures Production Targets are Met
- Keeps Cash Flow On-Plan



TERADATA

Other Time Based Use Cases

- Securities/stock trades
- Commodity prices tracking
- Autonomous trading algorithms
- Geospatial device tracking
- RFID/bar code packages
- Gaming – role play, gambling
- Mobile/web application event streams
- Predict staffing requirements
- Merchandise forecasting
- Infrastructure monitoring/audits
- Dev/Ops events
- Employee productivity and security
- Period based Sales analysis (Hourly sales)
- Transportation delays
- Vendor commitment attainment

Making Business Analytics Easier…

- Without time series capabilities
  - Complex SQL to align time intervals
  - Every query = full table scan
  - Complex analytics take hours
- With time series capabilities
  - Optimized data storage
  - Highly efficient queries plans
  - Reduce time to insight and action

TERADATA.

… and Data Scientists Happier

- Dramatic reduction in data preparation time
  - Its all ETL then ad hoc SQL
- Easiest way to join multi variate sensor streams
  - Align different grains of time
  - When you have to compare the shape of two curves
  - Reach across UDA with QueryGrid
- Teradata class scalability
  - Much higher accuracy
  - Turn development into deployment

TERADATA.

Teradata Database – Time Series Capabilities

Teradata Database 16.20

Agile Analysis enabled by Time-Aware Functions

- Time period aware aggregations
- Work with ANY time component data
- Impute missing values
    - Ignore, removed, update with constant

High-Performance enabled by Primary Time Index (PTI)

- Supports time sensitive decisions
- Fast access through:
    - Hash distribute by time bucket
    - AMP-local processing
    - Sequenced data

Time Aware Aggregate Example

```
SELECT $TD_TIMECODE_RANGE, $TD_GROUP_BY_TIME, SENSORID, AVG(TEMPERATURE) FROM BUOYS
    WHERE TIMECODE BETWEEN TIMESTAMP '2017-08-11 01:00:00' AND TIMESTAMP '2017-08-11 03:00:00'
    GROUP BY TIME( MINUTES(30) AND SENSORID)  USING TIMECODE(TD_TIMECODE)
    ORDER BY SENSORID, $TD_GROUP_BY_TIME;
```

Timecode-Range	Group by 30 minutes	Sensor ID	Temperature
'2017-08-11 01:00:00', '2017-08-11 **01:30:00**'	1	22	63.5
'2017-08-11 01:30:00', '2017-08-11 **02:00:00**'	2	22	64.6
'2017-08-11 02:00:00', '2017-08-11 **02:30:00**'	3	22	65.0
'2017-08-11 02:30:00', '2017-08-11 **03:00:00**'	4	22	65.1
,2017-08-11 01:00:00', '2017-08-11 **01:30:00**'	1	23	66.4
'2017-08-11 01:30:00', '2017-08-11 **02:00:00**'	2	23	65.1
'2017-08-11 02:00:00', '2017-08-11 **02:30:00**'	3	23	64.9
'2017-08-11 02:30:00', '2017-08-11 **03:00:00**'	4	23	65.1

Time Aware Aggregation Functions – GROUP BY TIME

Existing Aggregate Functions	
Average	Count
Describe	Kurtosis
Maximum	Minimum
Percentile	Rank
Skew	Sum
Std. population deviation	Std. sample deviation
Population variance	Sample variance

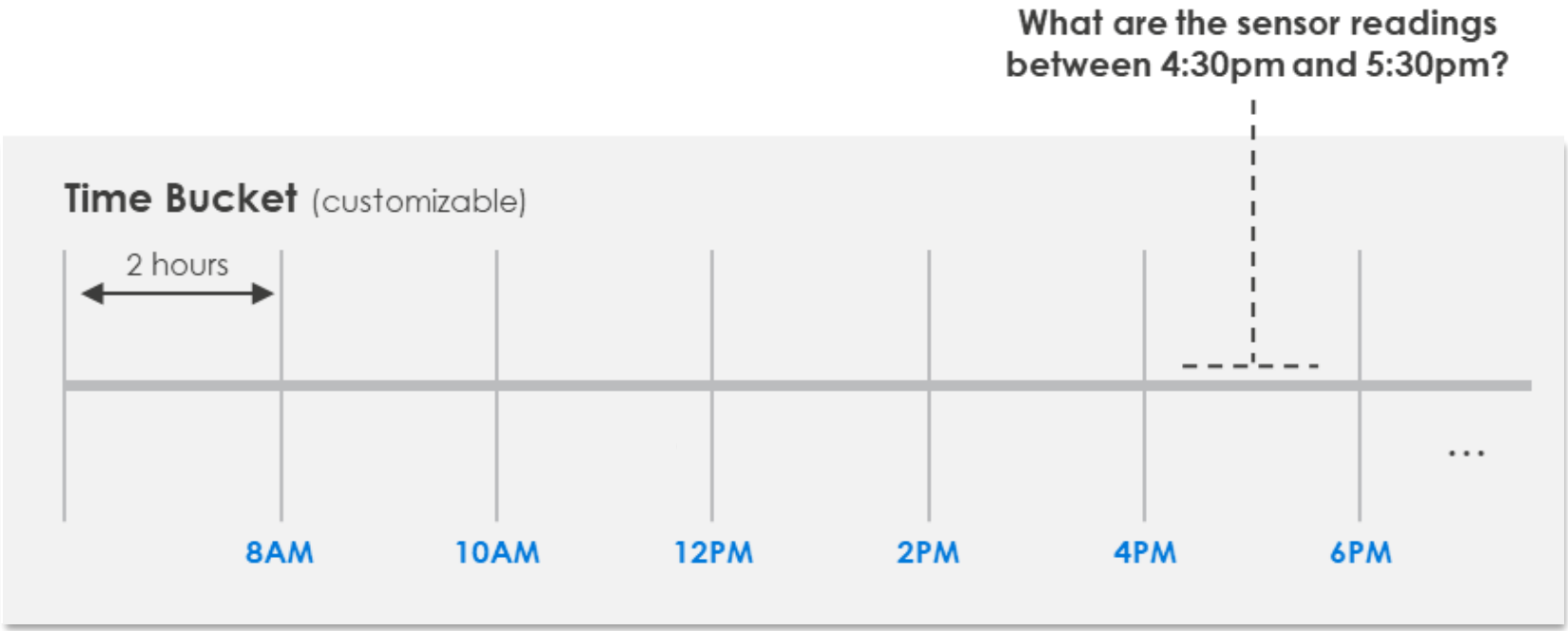
If not in the list above, then function is not time aware and cannot be used with the GROUP BY TIME clause

New Aggregate Functions	
Bottom	Delta_T
First	Last
Median	Mode
Top	Mean absolute deviation

These new aggregate functions are only invokable with the GROUP BY TIME clause

Primary Time Index (PTI)

High Performance Parallelism with Efficient Storage and Access



TERADATA.

Primary Time Index Tables (PTI)

Storage distribution choice	Time interval only	Time + column list	Column list only
	PRIMARY TIME INDEX (TIMESTAMP(6), DATE '2016-02-22', HOURS(2))	PRIMARY TIME INDEX (TIMESTAMP(6), DATE '2016-04-19', HOURS(2), COLUMNS(COUNTRYID,CARID))	PRIMARY TIME INDEX (TIMESTAMP(6), DATE '2016-01-01', COLUMNS(SENSORID))

In-table logical ordering	Time code only	Time code + sequence number

Many SQL Table Designs Include Time

	Partitioned Primary Index (PPI)	Temporal Tables	Primary Time Index (PTI)
Business	• Multi-dimensional analytics • Hierarchical analytics • Date, character, or numeric levels	• Time periods (ranges) • Historical relevance • Audit – what was the situation when…	• High volume time stamped data • Time aware analytics • Sorted data • Unique algorithms
Technology	• Multi-level (up to 64) • Does not effect row distribution to the AMPs • Data is not ordered	• Slowly changing dimensions • Insert, update, delete • Normalize and overlap functions	• Distribution to AMPs by time buckets • Updates/deletes rare • Insert late arrival data • Multivariate payload common

All table types can use "GROUP BY TIME"

Integrating Analytics Drive the Most Value

Assets

- Location data – TimeSeries and Geospatial
- Sensor observations – Timeseries

Open /Public Data

- Map data – Geospatial
- Traffic (Real Time; Historical) – JSON
- Emergency data – JSON
- Weather data – JSON / Temporal

Data Integration

Teradata Analytics Platform

Business Data (Transactional and Temporal):

- Requests
- Financial Data
- Historical data
- Supply Chain

Data Integration

Analytics

- Trip Planning
- Demand Based re-planning (Urgent Request bases)
- Optimal delivery plan
- Capacity prediction planning
- Repair / Replace decisions
- Delay versus Penalty costs

TERADATA.

Teradata Advantage

Integrated Analytics

- Analyze various data in context with business data
- Use multi-function analytic engine
- Leverage UDA ecosystem

High-performance enabled by **Primary Time Index (PTI)** that supports time sensitive decisions

…on **robust Teradata Database:** Scalable, highly available, high performance, and secure

Agile analysis enabled by **Time-Aware Aggregate Functions** that work with ANY time component data

… with **Teradata Everywhere** deployment options. Design for data gravity

TERADATA.

Key Takeaways Summary

- Integrate new data sources and include new user communities for broader insights
  - Examples that we shared such as incorporate IoT analytics and extending the tools into the data scientist arena for new insights

- Simplified and optimized capabilities to accelerate time based analytics
  - This would include the new time aware functions and optimized Primary Time Index table options

- Leverage integrated analytics with the scale, optimization, and management of Teradata
  - Examples that we shared such as the integration of Geospatial, Temporal, and Transactional to drive targeted and timely action


TERADATA®