

TAILORING ANALYTICS TO FIT A PANDEMIC RESPONSE

Delivering Analytics for COVID-19
Fast, Reliably, and for Everyone

Duncan Hamilton

Ryan Hawkins

Christian Minich



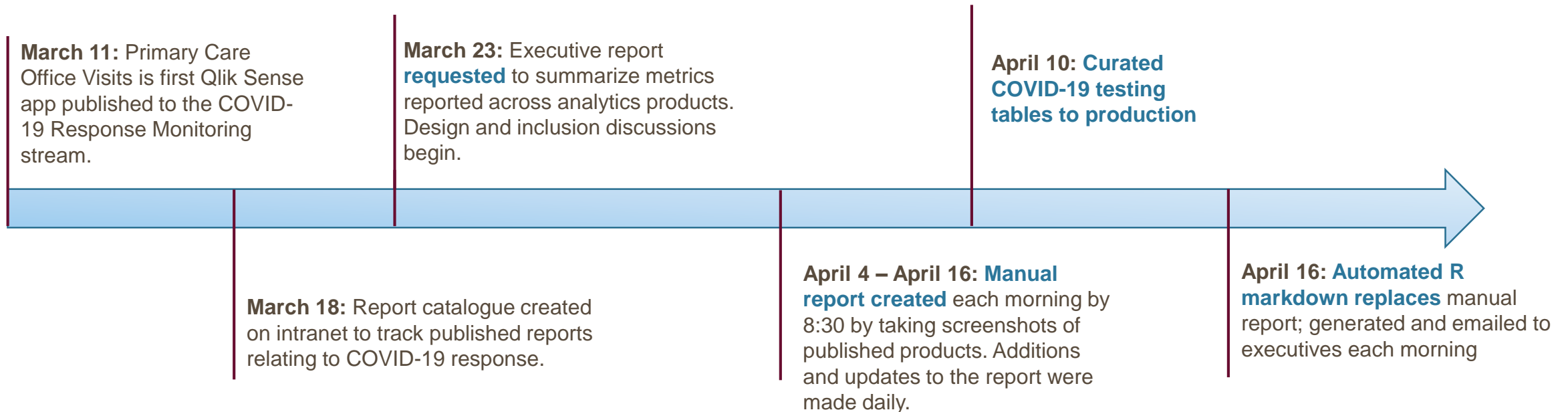
COVID-19 Presented New Data Challenges for CHOP

- A need for real-time and near-time data
 - Census, testing, employee exposures, supplies, etc
- Test results for everyone which required a different security approach
 - Patients, employees, DOH referrals, and others
- Suddenly everyone wanted data and wanted it blended
 - Clinical, operational, financial, supply chain, HR

The Response Demanded SPEED

- Accurate reporting within days
- Data to inform ramp-up plans in two weeks
- Accessible data to multiple parties and developers
- Infrastructure that was evolving on a daily basis that incorporated new processes and new validation tests

Timeline: From Start to Automation



In 36 days we went from the 1st data released to an automated Executive Report with concurrent contributions from 12 analysts

Our Data & Analytics Team as a Tailor Shop



THE TAILOR

Our Analysts have the domain experience and skillset to answer urgent, novel business questions



THE SHOP

Our analytics database (CHOP Data Blocks) is built to allow analysts to enrich it



THE CUSTOMERS

Our customers span the enterprise and have unique needs – clinical, finance, supply chain, HR

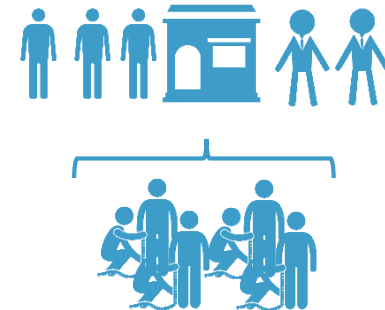
Part I: One Tailor, One Customer



Part II: Scale the Shop



Part III: 12 Tailors, Many Customers



<https://github.com/chop-analytics/presentations/blob/master/advancing-analytics-summer-seminar-2020/covid-19-data.pdf>

Part I: One Tailor, One Customer



- Analyst supports rapid go-live of a COVID Testing Team
 - CHOP Drive Through Testing sites opened on 3/18
 - Testing Patients, DOH referrals, and CHOP Employees
 - Employees referred through Occupational Health for testing; first time doing this at scale
 - Varying needs across Testing Team
 - Occ Health: time-sensitive communication with Employees for self-isolation/contact tracing/establishing return-to-work dates. Limited EPIC access
 - Drive Through Site Admin: Testing volumes over time, staffing needs
 - IP&C: Positive patients over time, positive rates
- How do we use our tools to quickly create a testing hub to support this needs of this team?

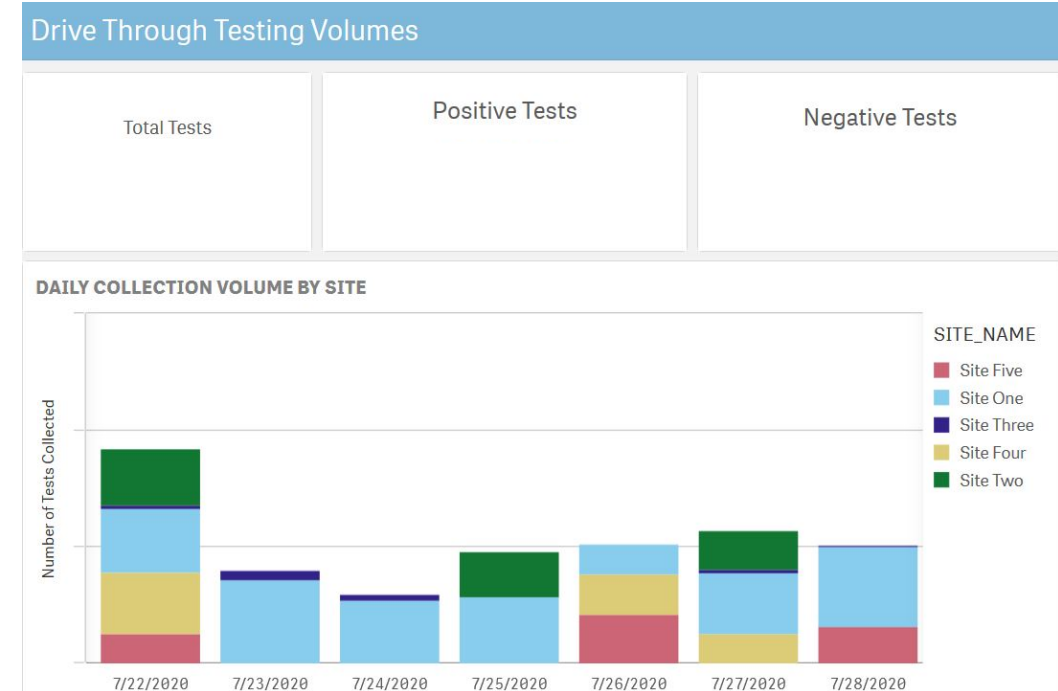
Analyst Uses Our Tools of the Trade

- CHOP Data Warehouse (CDW)
 - Curated analytical layer “Blocks” tables
 - Pathology team lab systems
- Newly-created REDCap Employee Intake survey
 - Used during referral to testing process
 - Second role as repository for Employment status info
- BI tool (Qlik Sense)
 - Work-streams with Data Governance support and ability to tailor access groups
- Working with stakeholders and SMEs
 - Interdisciplinary daily standup meetings
 - Quickly gathered evolving requirements and sourcing expertise from clinicians, process-focused advisors, members of pathology team

First Fitting: Success!



- Speed of service
 - Deployed Drive Through Testing Dashboard in 4 days
- Meeting immediate needs
 - Employee status and contact info
 - Testing volumes (Employee and Patient) by drive through site
 - Automated “push” positive/negative report sent daily to team for ease in contacting Employees
- Security and access
 - Established protected view of Employee and Patient info



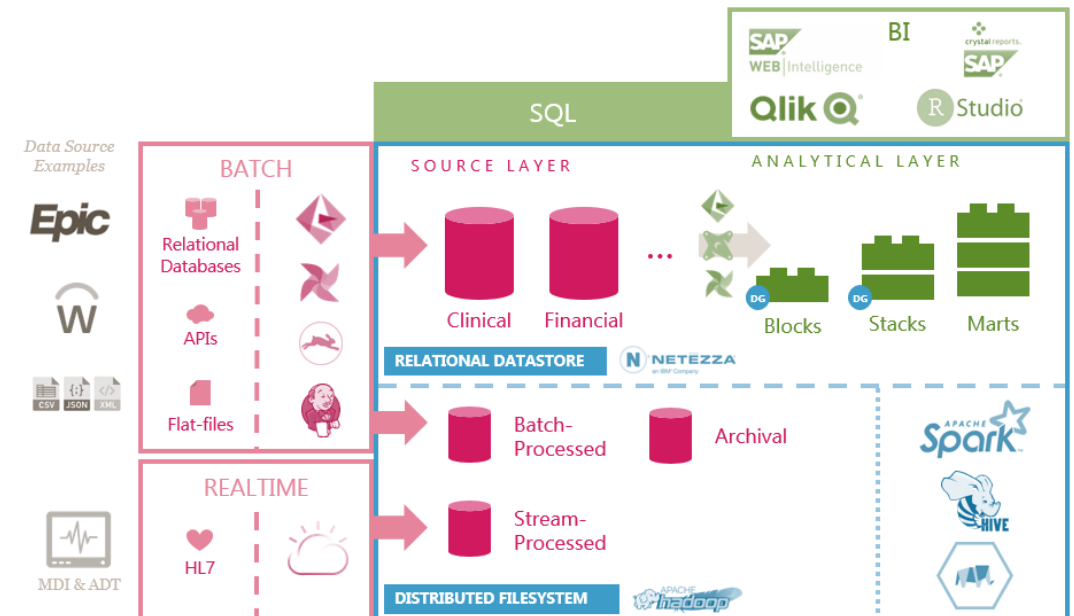
Testing Challenges: One Size Fits...One?



- Solution tailored to the use-case – not made for other customers
 - How can we take what we've built and answer other, related questions?
 - Inpatient testing?
 - Patient journey?
 - Equity?
 - CDC Person Under Investigation (PUI) Report form still manual
 - Logic validated and accessible, but not centralized
 - How can we be sure operational definitions remain aligned?
 - How do enact changes broadly/uniformly?
 - Security
 - How can we govern other uses of the same logic across the enterprise to ensure Patient and Employee information are protected?
- How do we spread this to other groups?

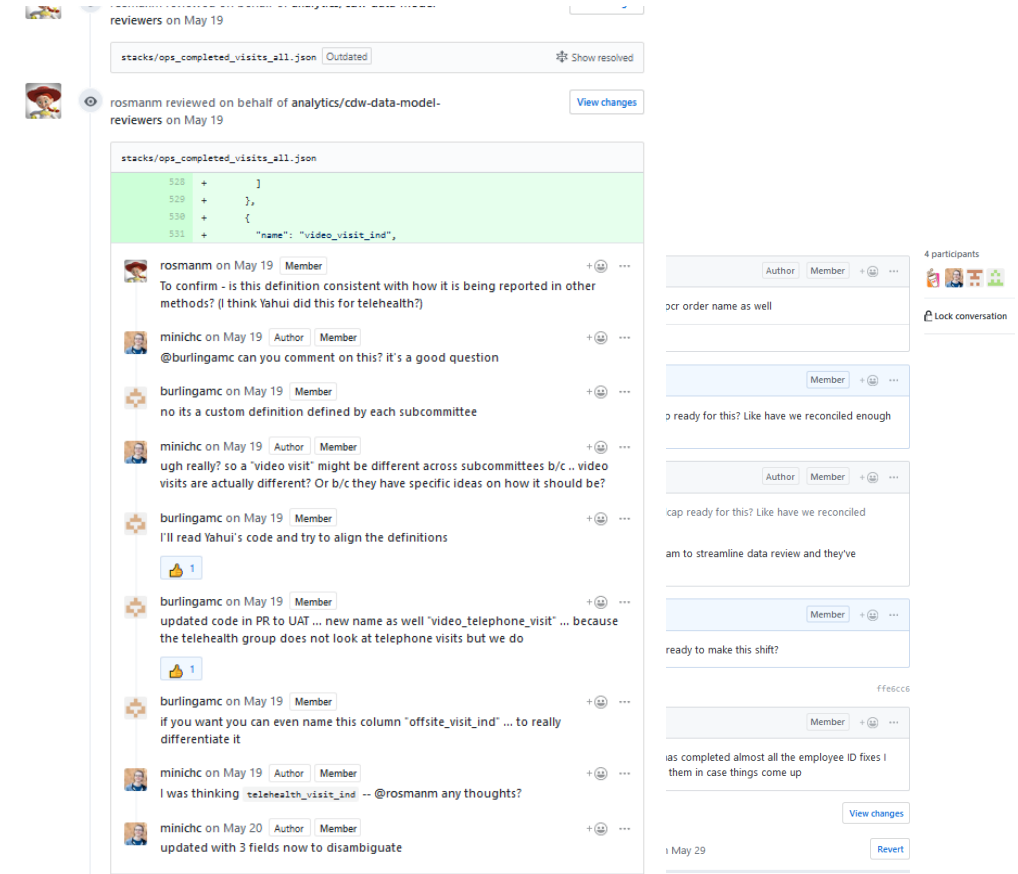
Part II: Scale the Shop

- Data in the enterprise data model should be
 - Governed in line with its visibility
 - Discoverable
 - Trusted
- Analysts close to the business should directly contribute to the enterprise data model



Governed in Line with its Visibility

- Definitions are reviewed by governance experts to ensure consistent definitions
- Contributors who are experts are consulted for major changes
- All SQL is code reviewed for SQL style and content



Data Should be Discoverable

- All data are documented using custom metadata-universe solution
- Approved definitions are viewable to the enterprise within our custom data dictionary application

The screenshot displays the 'METADATA UNIVERSE' application interface. At the top, there's a navigation bar with the logo and a user profile icon. Below it, a pink banner indicates '10 VALIDATION ERRORS'. The main content area shows the details for the 'stacks.outbreak_covid_cohort_patient' table. It includes a description, load type ('truncate_reload'), security groups ('chop_blocks_clinical_data'), and a list of columns with their descriptions and source counts. A sidebar on the right contains links for 'EDGE Admin', 'Report an Issue', and 'Download'. At the bottom, there's a section for 'CDW TABLE' with a filter column list and a table of columns and their descriptions.

stacks.outbreak_covid_cohort_patient

DESCRIPTION: This table provides information on the CHOP Patients of individuals being testing for COVID-19 SARS coronavirus 2 RNA. It does not include CHOP employees or other adults who have been tested

LOAD TYPE: truncate_reload

SECURITY GROUPS: chop_blocks_clinical_data

COLUMNS:

NAME	DESCRIPTION	SOURCES
* 123 proc_ord_key	The key generated by the CDW that is associated ...	1
ABC patient_name	The first and last name of the patient as entered in...	1
ABC last_nm	The last name of the patient	1

Back to Search Results

CDW TABLE

outbreak_covid_cohort_patient

This table provides information on the CHOP Patients of individuals being testing for COVID-19 SARS coronavirus 2 RNA. It does not include CHOP employees or other adults who have been tested

Type: not applicable

Database: stacks (chop_analytics)

Source Tables: procedure_order_clinical, outbreak_covid_cohort

Filter column list

Columns	Description	EPIC INI Items
proc_ord_key	The key generated by the CDW that is associated with the procedure order	
abnormal_result_ind	Indicates whether the order was flagged as being abnormal -- in this case, positive for COVID-19	

Data Should be Trusted

- All data are validated using a testing framework in the data build tool (dbt)
 - Assertions include: not null; primary key; is an indicator; date in past; referential integrity
- Test results are run after each ETL run, stored in our database, and exposed within a BI Application

Branch: master ▾ stacks-outbreak / etl / outbreak_covid_cohort_patient_encounter.yml

☐ Minich update tests for changes

0 contributors

29 lines (29 sloc) | 699 Bytes

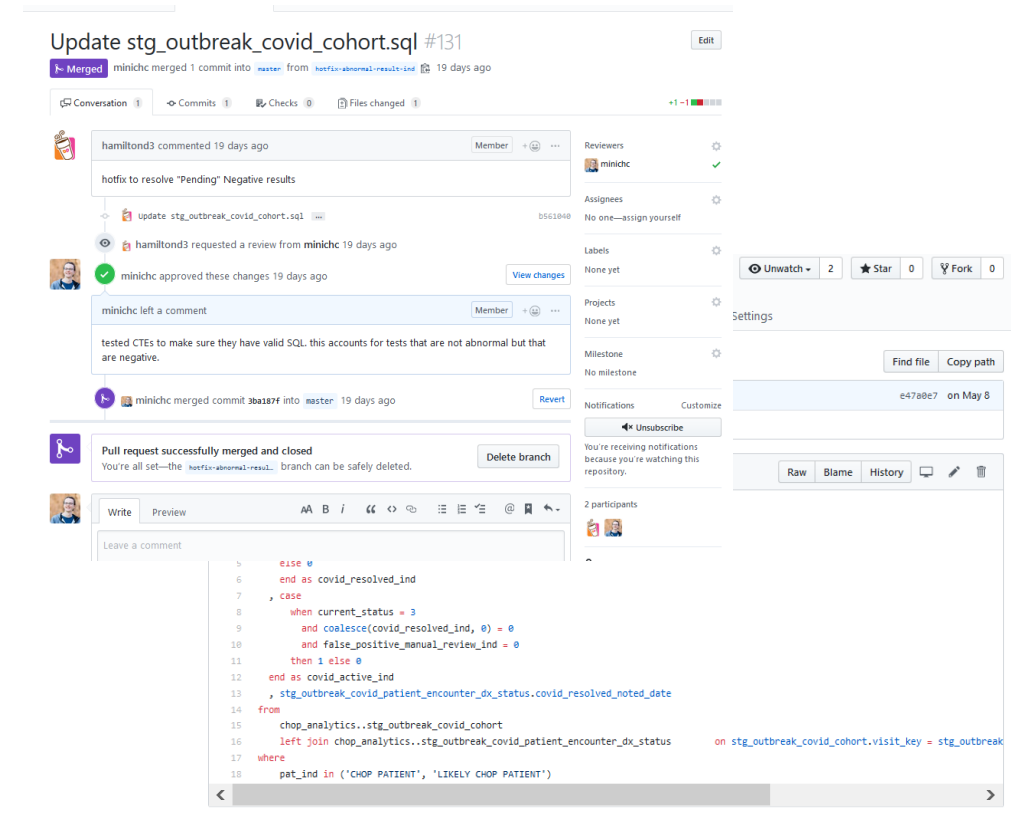
```
1 version: 2
2 models:
3   - name: outbreak_covid_cohort_patient_encounter
4     tests:
5       - dbt_chop_utils.is_primary_key:
6         column_names: [VISIT_KEY]
7     columns:
8       - name: VISIT_KEY
9         tests:
10          - not_null
11       - name: PATIENT_NAME
12       - name: MRN
13         tests:
14          - not_null
15       - name: CSN
16         tests:
17          - not_null
18       - name: ENCOUNTER_DATE
19         tests:
20          - not_null
```

dbt Test Results

Table	Column	Test Type	Date	Failed?
adt_bed	bed_name	not_null	8/16/2020	No
adt_bed	currently_admitted_ind	is_indicator	8/16/2020	No
adt_bed	department_group_name	not_null	8/16/2020	No
adt_bed	exit_date	is_greater_or_equal_to	8/16/2020	Yes
adt_bed	last_bed_ind	is_indicator	8/16/2020	No
adt_bed	-	is_primary_key	8/16/2020	No
adt_bed	-	recency	8/16/2020	No
adt_department	currently_admitted_ind	is_indicator	8/16/2020	No
adt_department	hsp_acct_key	fill_rate	8/16/2020	No
adt_department	last_department_ind	is_indicator	8/16/2020	No
adt_department	-	is_primary_key	8/16/2020	No
adt_department_group	currently_admitted_ind	is_indicator	8/16/2020	No
adt_department_group	last_department_group_ind	is_indicator	8/16/2020	No

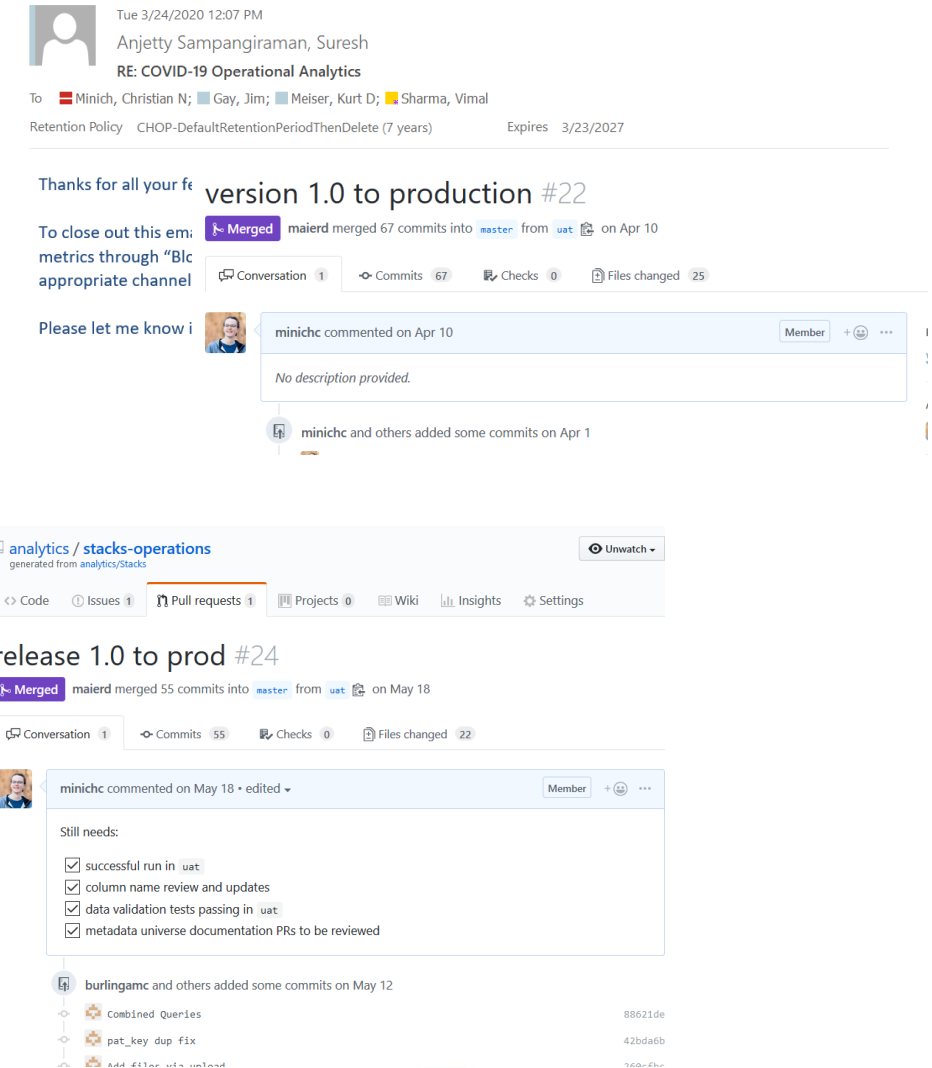
Analysts Contribute Directly to the Data Model

- All code is hosted on GitHub (version control system)
 - Analysts have access to push changes to development and user acceptance testing
- Hosted a semi-weekly standup with all analyst contributing to discuss work and business needs



But Can it be Done Fast?

- COVID testing data mart
 - First release: 18 days, 4 tables
 - Test results for employees and patients
 - Regulatory reporting needs for DHS
 - Employee exposure survey data
- Contact tracing data mart
 - First release: 9 days, 1 table
 - Patient – provider interactions
 - Patient – patient interactions
- Completed and Scheduled Visits data mart
 - First release: 27 days, 4 tables
 - All scheduled encounters
 - All completed encounters



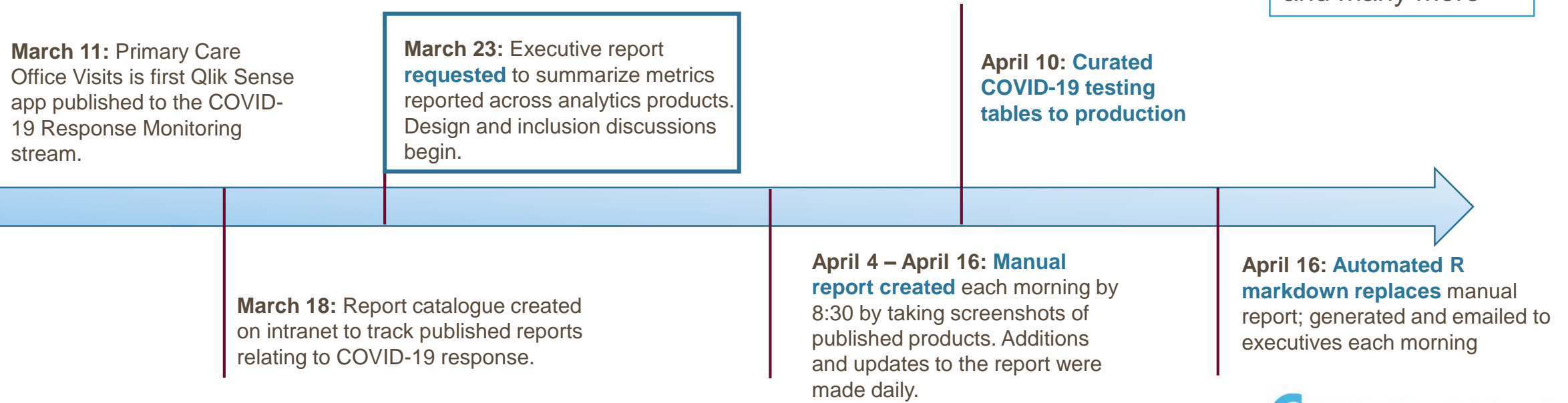
Part III: 12 Tailors, Many Customers



Stitching multiple customer needs together to provide situational awareness across the enterprise

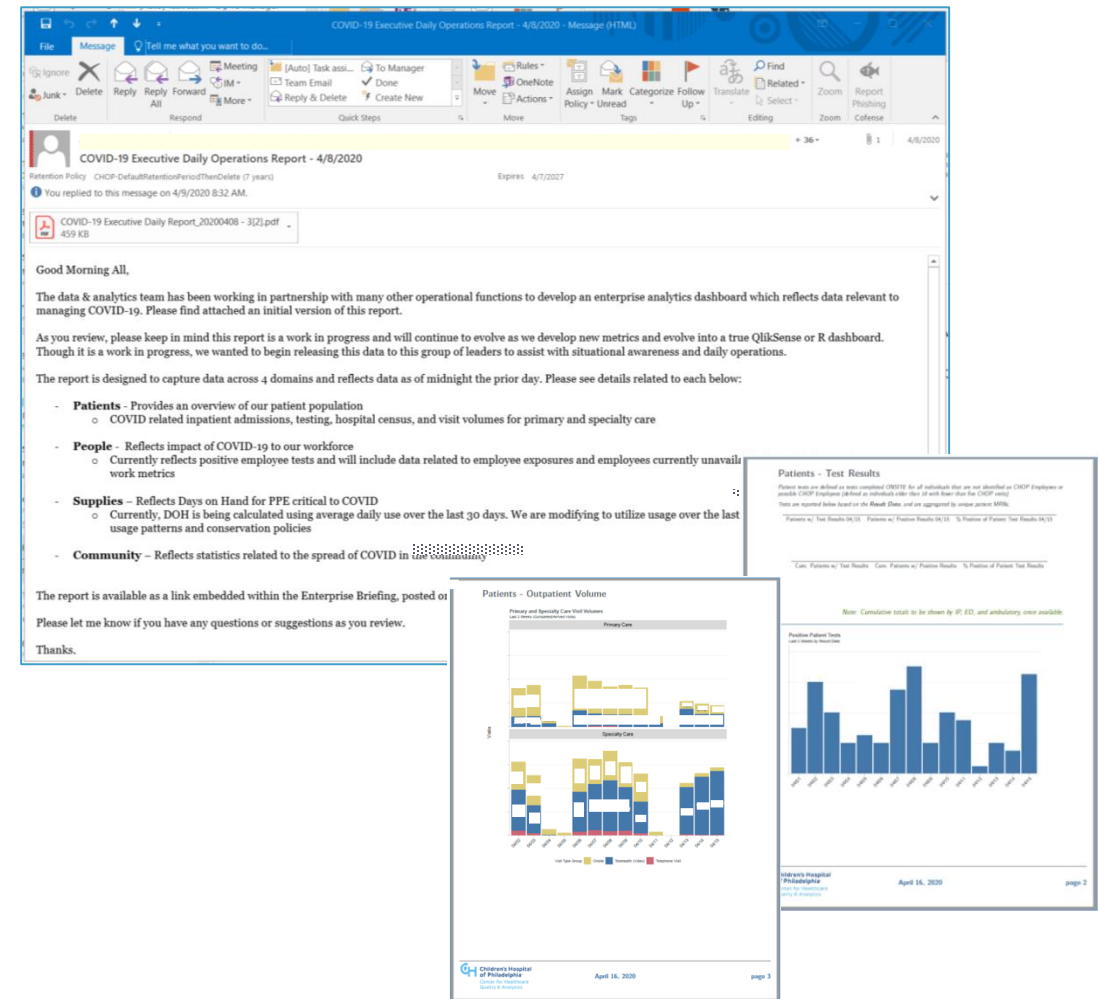
Customers:

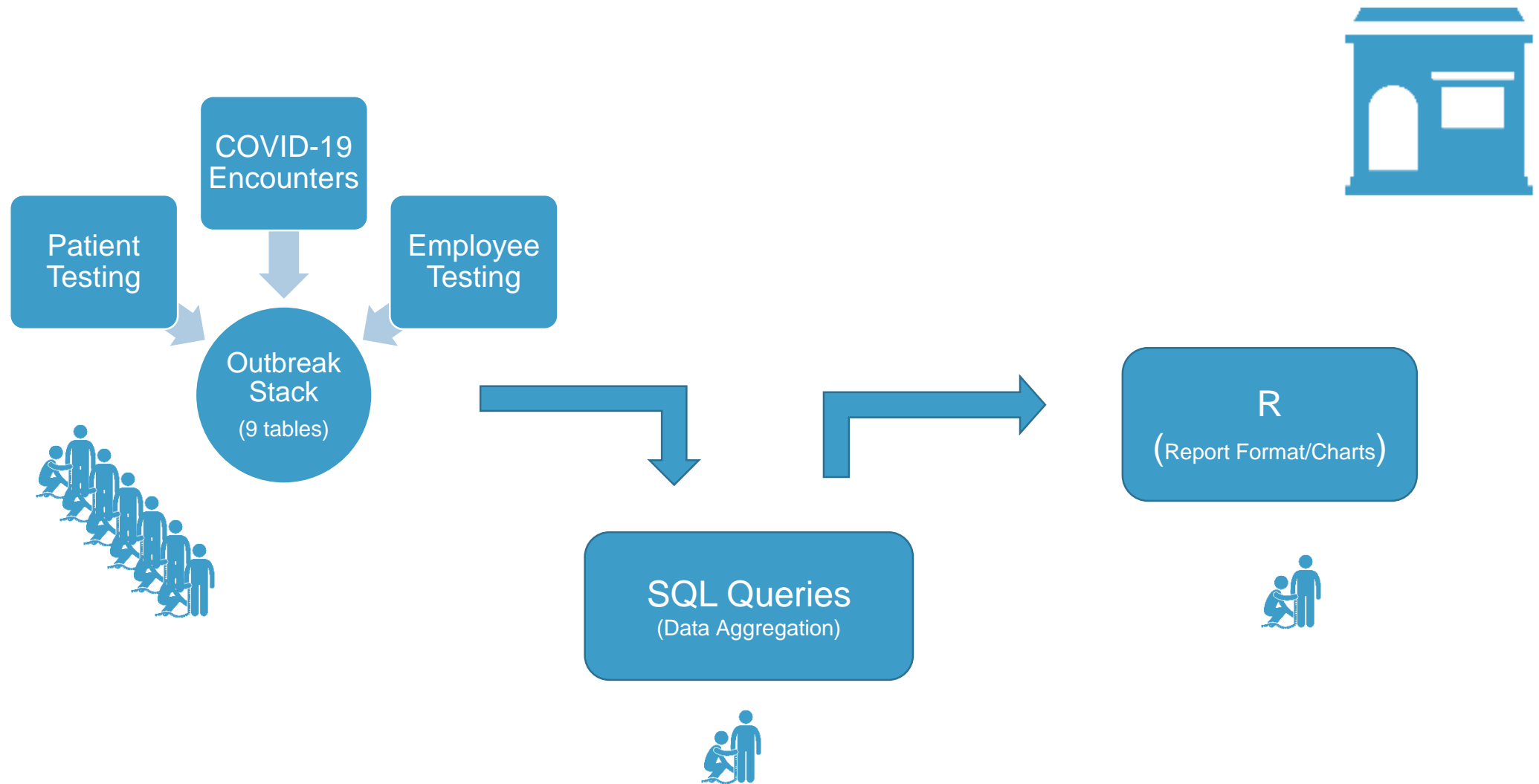
Executives
Operations
Finance
Supply Chain
HR
Research
and many more



One Enterprise View

- Executive summary – KPIs and trends, not details
- Combined data for **patients, staff, supplies**, and surrounding **community**
- 100% alignment with other analytics products (reports, dashboards, tools)
- Sent daily via email and published on intranet





Phase 1: Testing and Exposure

- Are infections being controlled in the hospital?
- How is the lockdown affecting outpatient volume?
- How many:
 - COVID-19 patients are currently admitted
 - Tests were given yesterday/How many positive
 - Outpatient visits are turning into telehealth visits
 - PPE Supplies do we have on hand

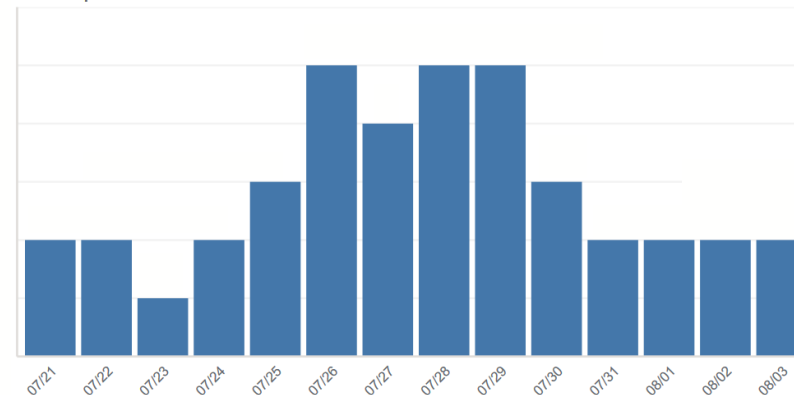
Yesterday's Test Results 08/03

Patients w/ Test Results	Patients w/ Positive Results	% Positive of Patient Test Results
		%

Cumulative Test Results

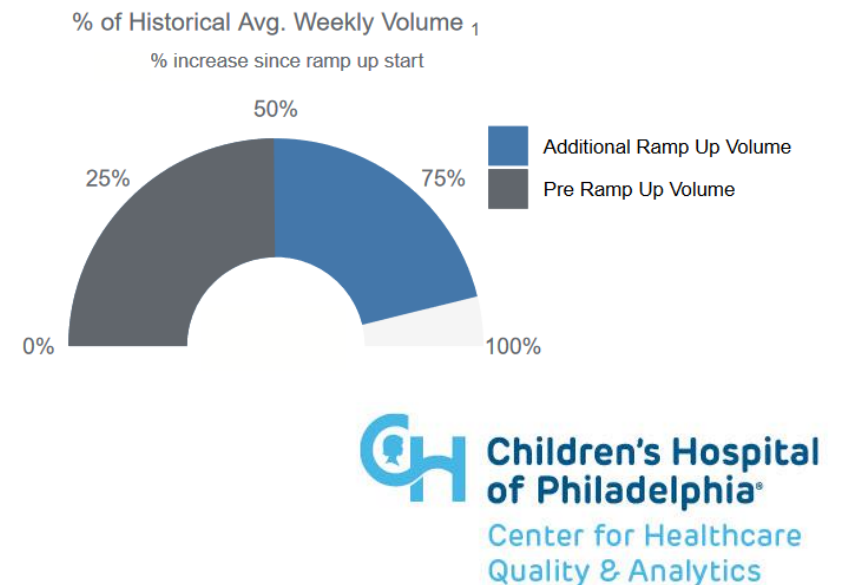
Patients w/ Test Results	Cum. Patients w/ Positive Results	% Positive of Patient Test Results
		%

COVID Inpatient Census Over Last 2 Weeks



Phase 2: Ramp Up

- How are volumes compared to last year at this time?
- How are volumes compared to February 2020?
- Are we doing better than last week?
- Which service area are recovering, which are struggling?
- Weekly updates from each service area?



Phase 3: Financial Impact

- What does the new normal look like?
- What financial contingencies are needed?

Financial Performance

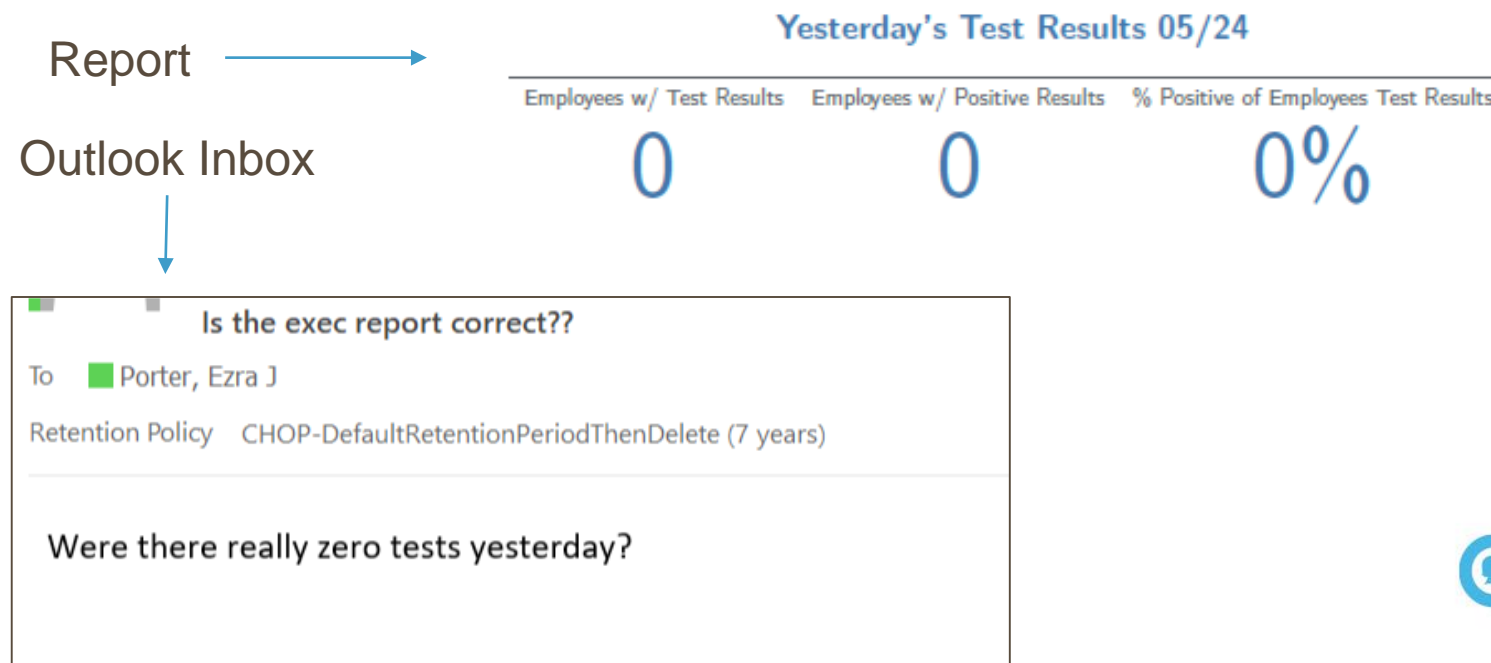
In addition to the summary of nursing unit occupancy trends below, detailed information related to hospital's gross revenue and financial statistics is available at the [Finance Daily Dashboard](#). This tool allows users to track and trend gross revenue and statistics by area and department on a daily, monthly, and yearly basis. Users can drill down to various levels of granularity including inpatient/outpatient and cost center.

Key Finance Terms and Definitions:

- Variance - Actual results compared to budget
- Total Gross Revenue - Hospital IP/OP, Care Network and Home Care posted charges
- Average Daily Census - Total patient days divided by number of days in the time period
- Adjusted Patient Days - Inpatient days adjusted by the outpatient adjustment factor (IP Days x (OP gross revenue/IP gross revenue))

Challenges: Scheduling

- Report emailed to ~120 hospital leaders at 7am and URL shared in the daily operations briefing
- Large amount of near real-time data



Successes: Changing Content

- Requirements have changed as pandemic evolved
- “How are we doing today?” → “How does this month compare to last?”
- Report infrastructure supports painless(-ish) additions of new displays...
 - Ramp-up volumes, rapid test inventory, financial performance
- and collaboration between analysts on crafting additions

Conclusion

- COVID-19 is once-in-a-lifetime, the analytics are not
 - One Tailor, One Customer: Experience + skillset to deliver exactly what the customer needs
 - Scale the Shop: Curated data foundation and process to scale the tailor
 - 12 Tailors, Many Customers: Using the materials of the shop so that many tailors can collaborate on many data product that span the business
- We received widespread adoption of our COVID-19 data infrastructure
 - **41** active users of our tables (direct SQL query writers)
 - **75** avg. daily active users of the COVID executive summary (1,600 views in the past month)
- It takes a team
 - Analysts: **12** analysts contributed to the COVID data infrastructure; **13** created the deliverables used
 - Platform engineers: **3** software engineers who developed the tech stack that powers the curated data layer and process to enrich it
 - Platform operations: **4** ops engineers who maintain the platform

Appendix

CHOP Data Warehouse

Data Source Examples

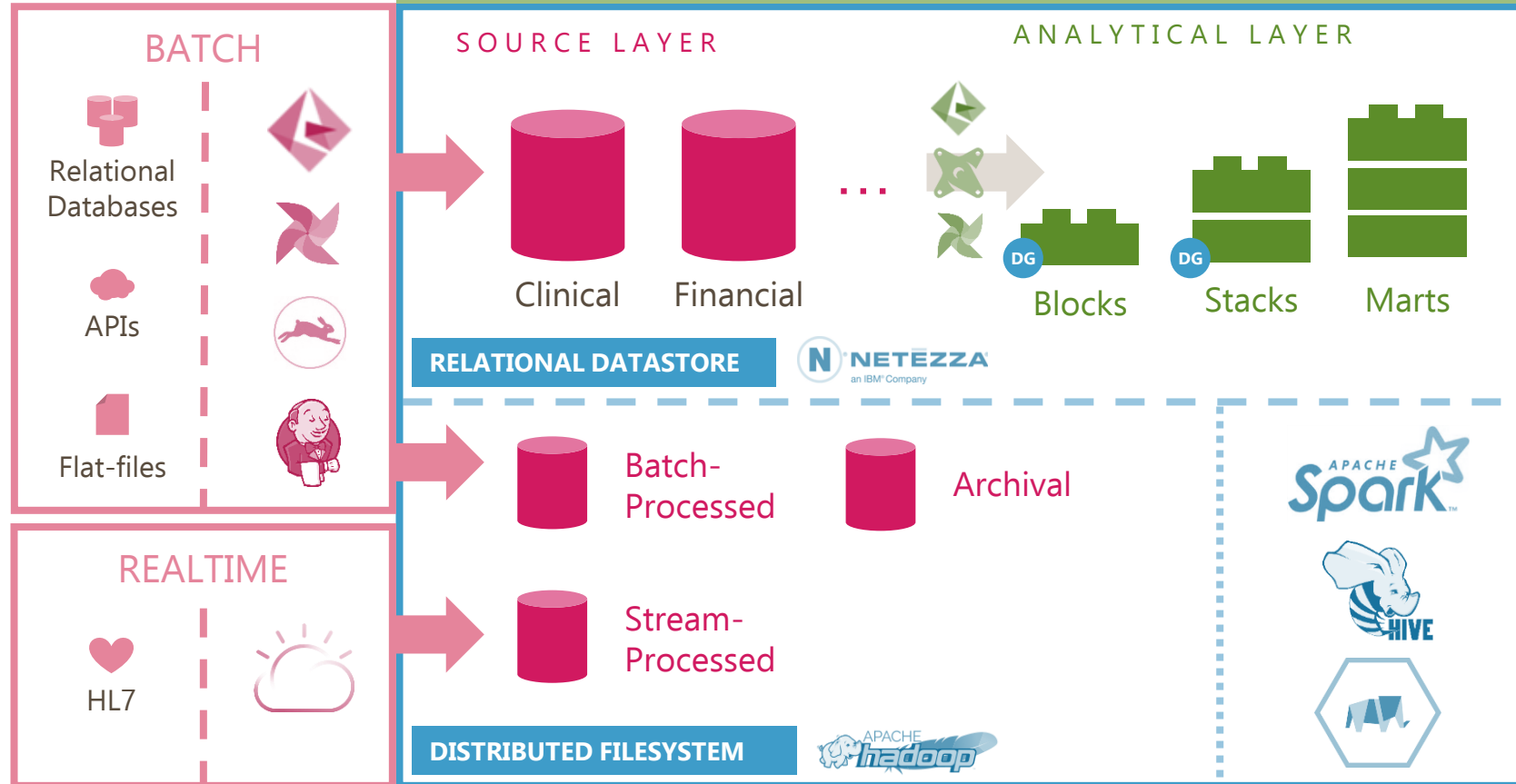
Epic

W

CSV JSON XML



MDI & ADT



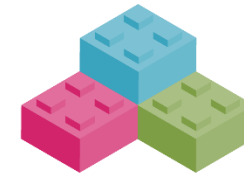
CHOP Blocks

Semantic layer for CDW to standardize definitions

- Make routine querying simpler
- Maintain consistency in data, metrics, and definitions
- Serve as a foundation for self service
- Be Use Case Driven (80%)

The blocks currently have tables relating to

- Encounters
- Flowsheets
- Medications
- Diagnoses
- Procedure orders
- Surgical procedures
- Patient flow



CHOP
BLOCKS

CHOP Data Stacks

CHOP Data Stacks are content area owned, governed set of tables that sit on top of base CHOP Data Warehouse (CDW).

Use Cases

- Governed data marts that should be shared broadly and power self-service for content area experts (e.g. capacity, neonatology)
- Governed data marts that are shared across teams (e.g. data for the frontier programs)

Purpose:

- Developed (generally by analysts) in specific domain areas (Harm, Ortho, etc.) to standardize the data they frequently use
- Shared with multiple groups (external to domain) across the enterprise
- Meant to be *mostly* permanent pieces of data infrastructure (not short term projects).
- Fully governed, the definitions are reviewed and agreed upon by both domain experts and Data Governance.

- Frictionless standard
- GitHub-integrated editor



R-Studio Connect

- Data Science Publishing Platform
 - Scheduled static documents
 - Interactive dashboards
 - APIs
 - Secured document access using LDAP
- Primary User Groups:
 - Center for Healthcare Quality and Improvement
 - Data Engineering
 - Data Analytics
 - Pathology

