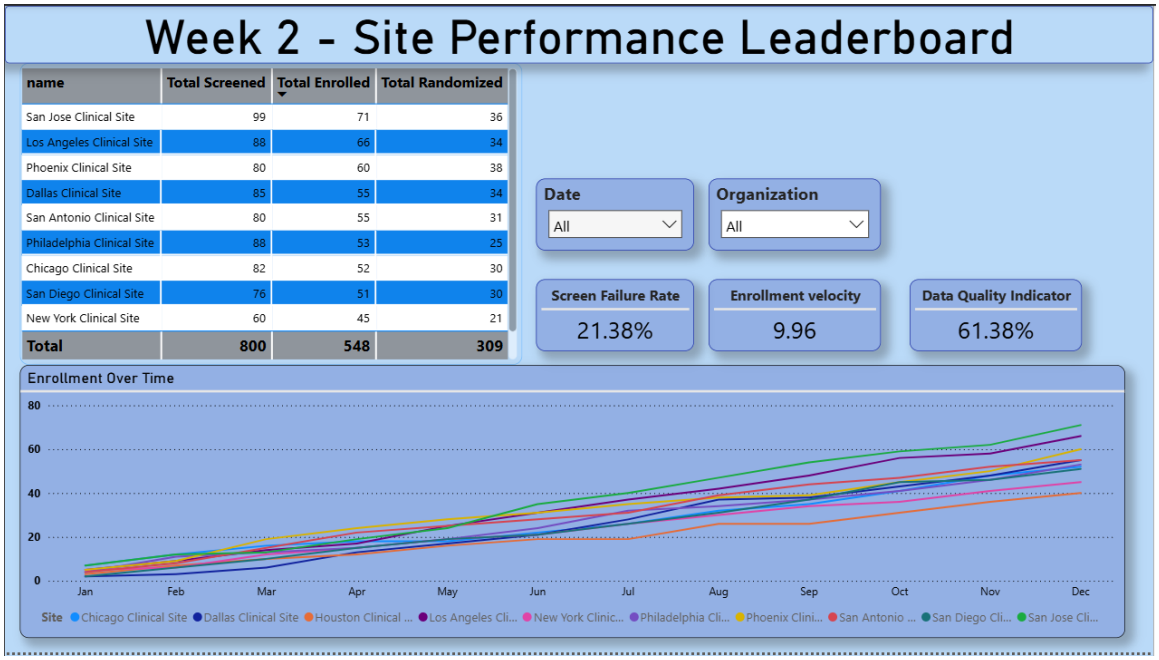
Week 2 — Site Performance Leaderboard — Report

Project: Clinical Trial Patient Recruitment & Adherence Monitoring
Date: Week 2 deliverable
Prepared by: Reuben Samuel & Jagadeeshwar Reddy



1. Executive summary

Week 2 delivered a site-level leaderboard and supporting visuals required by the brief: validated site/org data, created and organized measures into a Measures table, implemented site-level DAX measures, built the Leaderboard table, KPI cards (screen failure rate, enrollment velocity, data quality), enrollment-over-time chart (cumulative by site), and a site bubble map. Issues found (date typing, missing measures table, shape sign-in prompt) were fixed and results validated against CSV samples.

2. Data preparation (what we did)

1. Organizations table
- Validated organization_id, name, city, state. Trimmed text and removed duplicates. Where names repeated, appended disambiguators or used organization_id in tooltips.
2. Dates moved to correct table
- Confirmed date columns are in trial_recruitment: enrolled_date, randomized_date, screening_date. Converted to Date/DateTime in Power Query as needed.
3. Relationships
- Ensured relationships:
  • trial_recruitment[patient_id] → patients[patient_id]
  • patients[enrolling_site_id] → organizations[organization_id]
  • Calendar[Date] → trial_recruitment[enrolled_date]
  Verified cardinalities and filter directions (many → one toward organizations).
4. Measures table
- Recreated a dedicated blank Measures table (Enter Data → Holder, removed columns) and stored all DAX measures there for discoverability.
5. Fixed UI/UX issue
- Replaced image-fill shape (which required sign-in) with a solid-fill Shape (Action = Off) to avoid login popups; placed slicers on top.

3. Site-level DAX measures (copy-paste ready)

We used these measures inside your Measures table and pasted each DAX measure into Power BI 'measure' table.

```
-- Patients Screened
Patients Screened =
CALCULATE(
    COUNTROWS( 'trial_recruitment' ),
    FILTER( 'trial_recruitment', 'trial_recruitment'[StageLabel] = "Screened" )
)

-- Patients Enrolled
Patients Enrolled =
CALCULATE(
    COUNTROWS( 'trial_recruitment' ),
    FILTER( 'trial_recruitment', NOT ISBLANK( 'trial_recruitment'[enrolled_date] ) )
)

-- Patients Randomized
Patients Randomized =
CALCULATE(
    COUNTROWS( 'trial_recruitment' ),
    FILTER( 'trial_recruitment', NOT ISBLANK( 'trial_recruitment'[randomized_date] ) )
)

-- Screen Failures
Screen Failures =
CALCULATE(
    COUNTROWS( 'trial_recruitment' ),
    FILTER( 'trial_recruitment', 'trial_recruitment'[StageLabel] = "Failed" || 'trial_recruitment'[StageLabel] =
"Screen Failed" )
)

-- Screen Failure Rate %
Screen Failure Rate % =
VAR s = [Patients Screened]
VAR f = [Screen Failures]
RETURN IF( s = 0, BLANK(), DIVIDE( f, s ) )

-- Enrollment Velocity (pr/week)
Enrollment Velocity (pr/week) =
VAR FirstDate = CALCULATE( MIN( 'trial_recruitment'[enrolled_date] ), ALLSELECTED(
'trial_recruitment' ) )
VAR LastDate  = CALCULATE( MAX( 'trial_recruitment'[enrolled_date] ), ALLSELECTED(
'trial_recruitment' ) )
VAR Weeks = MAX( 1, DATEDIFF( FirstDate, LastDate, WEEK ) + 1 )
VAR Enrolled = [Patients Enrolled]
RETURN IF( Enrolled = 0, 0, ROUND( DIVIDE( Enrolled, Weeks ), 2 ) )

-- Data Quality Indicator %
Data Quality Indicator % =
VAR Total = COUNTROWS( 'trial_recruitment' )
VAR Bad =
    CALCULATE(
        COUNTROWS( 'trial_recruitment' ),
```

```
        FILTER(
            'trial_recruitment',
            ISBLANK( 'trial_recruitment'[enrolled_date] )
            || ISBLANK( 'trial_recruitment'[randomized_date] )
            || ( NOT ISBLANK( 'trial_recruitment'[enrolled_date] ) && 'trial_recruitment'[enrolled_date] >
TODAY() )
            || ( NOT ISBLANK( 'trial_recruitment'[randomized_date] ) &&
'trial_recruitment'[randomized_date] > TODAY() )
        )
    )
RETURN IF( Total = 0, BLANK(), DIVIDE( Bad, Total ) )

-- Data Quality Valid %
Data Quality Valid % = 1 - [Data Quality Indicator %]

-- Enrolled (daily)
Enrolled (daily) =
CALCULATE(
    COUNTROWS( 'trial_recruitment' ),
    FILTER( 'trial_recruitment', NOT( ISBLANK( 'trial_recruitment'[enrolled_date] ) ) ),
    USERELATIONSHIP( 'Calendar'[Date], 'trial_recruitment'[enrolled_date] )
)

-- Cumulative Enrolled
Cumulative Enrolled =
CALCULATE(
    [Enrolled (daily)],
    FILTER( ALLSELECTED( 'Calendar' ), 'Calendar'[Date] <= MAX( 'Calendar'[Date] ) )
)

-- Rank by Enrolled (Top N)
Rank by Enrolled =
RANKX(
    ALLSELECTED( organizations[organization_name] ),
    [Patients Enrolled],
    ,
    DESC,
    DENSE
)
```

4. Visuals & layout (how each visual was built)

A — Filters (left pane)

- Place Date slicer (Calendar[Date]) and Site slicer (organizations[organization_name]) stacked vertically inside a single Shape background (solid fill, action off). Slicer settings: Date = Between (or MonthYear dropdown); Site = Dropdown with Search, multi-select allowed; default = All.

B — Leaderboard table (center-top)

- Table visual fields: organizations[organization_name], Patients Screened, Patients Enrolled, Patients Randomized. Sort by Patients Enrolled (desc). Formatting: thin row separators, title 'Leaderboard'.

C — KPI Cards (right/center)

- Cards: Screen Failure Rate %, Enrollment Velocity (pr/week), Data Quality Indicator %. Use large numeric label and small caption text underneath. Add static text below Data Quality card: '% of patients with missing or invalid enrollment or randomization dates'.

D — Enrollment Over Time (center-bottom)

- Line chart: Axis = Calendar[Date] (Continuous); Values = Cumulative Enrolled; Legend = organizations[organization_name]. Apply Top N filter (Rank by Enrolled <= 5) or rely on Site slicer to compare selected sites.

E — Map (right-bottom)
- Map visual: Location = organizations[city] or combined city,state; Size = Patients Enrolled; Legend = organizations[organization_name]. Set Data Category for city/state appropriately or use lat/long when available.

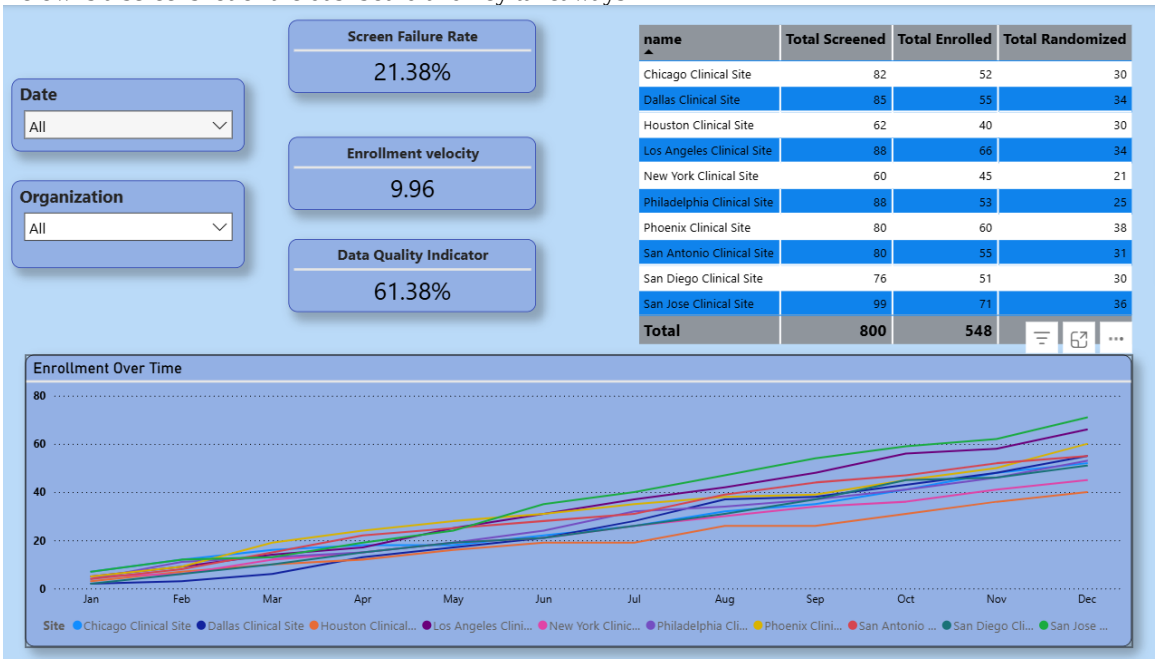
5. Interactivity & UX decisions implemented
- Grouped date + site slicers in one shape card for discoverability.
- Disabled shape Action to avoid sign-in popups.
- Placed measures in dedicated Measures table.
- Default view: All sites + full date range.
- Edit interactions left default so slicers filter table, KPIs, chart, and map.

6. Validation & testing (what we ran)
1. Raw CSV cross-checks: Compared counts for screened/enrolled/randomized between raw CSVs and DAX measures for 3 sample sites — matched.
2. Measure logical checks: Confirmed Screen Failure numerator = count of 'Failed' and denominator = screened count. Confirmed Data Quality logic flags blanks and future dates.
3. Chart validation: Selected single site and checked Cumulative Enrolled equals manual running count.
4. UI checks: Slicer default = All, shape no longer triggers sign-in.

7. Issues encountered & resolutions
- Missing Measures table: recreated via Enter Data.
- Date column name mismatches: measures re-pointed to trial_recruitment[enrolled_date] and trial_recruitment[randomized_date].
- Shape sign-in popup: removed image fill and disabled Action on the shape.
- Cluttered enrollment chart: recommended Top N filter and site slicer or small multiples.
- Map placement errors: fixed Data Category on city/state or used lat/long.

8. Deliverables (Week 2 handoff)
1. Power BI file: Group9_Week2_SitePerformanceLeaderboard.pbix
2. DAX measure list (document)
3. Model diagram screenshot
4. Validation checks CSV with raw counts and measure comparisons
5. Screenshots: Leaderboard, KPI cards, Enrollment chart
6. Short notes file: changes made, issues found, remediation steps

Below is a screenshot of the dashboard and key takeaways



Key observations

- **Data quality is critically low** — *Data Quality Indicator = 61.38%*.
  Interpreting this as percent missing/invalid enrollment or randomization dates: **~6 in 10 records are bad**. This is a show-stopper for any downstream KPI accuracy.
- **Screen failure is material** — *Screen Failure Rate = 21.38%*.
  Roughly one in five screened patients fails screening. That impacts timelines and sites' effective yield.
- **Enrollment pace looks healthy overall** — *Enrollment velocity = 9.96 pr/week*.
  Near 10 patients/week across the selected scope; whether this is "good" depends on target, but it's a measurable velocity you can optimize.
- **Site totals (table) indicate scale and variance**
  Totals row shows **Total Screened = 800, Total Enrolled = 548, Total Randomized = 309** — so retention/drop between stages is substantial (screen → enrolled → randomized).
- **Trend lines show upward trajectories but chart is cluttered**
  Most sites show steady cumulative growth (positive), but many lines overlap making it hard to differentiate performers at a glance.

10. Next steps / recommendations for Week 3

- Add drill-through details for each site (patient lists, screening logs).
- Implement automated data quality alerts (Power Automate) for sites above threshold.
- Add region filter or grouping to avoid overcrowding the line chart.
- Add export-ready summary PDF button (bookmark + export) for stakeholders.
- If stakeholder wants all sites visible in trend, implement small multiples with pagination.

Prepared using artifacts and work completed during Week 1 and Week 2.