



Getting Your Data Ready for BI

Preparing data is the single biggest factor in BI success.

8

Checklist
Categories

40+

Action
Items

5

Quality
Gates



Sourcing & Permissions

- Identify all data sources
- Check API limits & extraction constraints
- Confirm access credentials & permissions
- Define data ownership & stewardship
- Verify data refresh/update frequency



Clean, Consistent, Correct

- Check for missing / null values
- Remove duplicates
- Standardize date/time formats
- Standardize currencies & units
- Validate key fields (IDs, emails, SKUs)
- Fix obvious inconsistencies
- Confirm totals match source systems



Schema & Relationships

- Define business questions & KPIs first
- Establish primary keys
- Identify required dimensions
- Establish relationships / joins
- Identify required measures
- Decide star vs snowflake schema
- Create consistent naming conventions
- Avoid many-to-many relationships



Shaping Data for Analysis

- Clean & normalize values
- Handle outliers
- Create calculated fields
- Create derived metrics
- Aggregate where appropriate
- Document all transformations
- Split combined fields (full name → first/last)



Protect, Define, Comply

- Define data definitions
- Confirm compliance (GDPR, HIPAA, etc.)
- Apply role-based access controls
- Document metric logic
- Mask sensitive data if needed



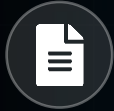
Speed & Efficiency

- Remove unused fields
- Limit row counts where possible
- Index key columns
- Validate query speed
- Pre-aggregate large datasets



Verify Before You Publish

- Cross-check with source reports
- Confirm totals & subtotals
- Test edge cases
- Test refresh schedules
- Verify filters behave correctly



Record Everything

- Data source inventory
- Known limitations
- KPI definitions
- Data refresh schedule
- Transformation rules

DATA READINESS IMPACT

**Clean data
accelerates BI
time-to-insight
significantly**

100%

Before Data Prep

After Checklist

faster insights

Your data is ready when it passes all five gates.



Accurate

Numbers match source



Consistent

Same logic everywhere



Timely

Refreshes on schedule



Secure

Access controlled



Actionable

KPIs measurable

