Table of Contents

# 📚 Video Analytics Project - Navigation Guide

**Your Role-Based Guide to Video Analytics Documentation**

Welcome! This repository is organized to support every role in the Scrum team - from executives to developers. Find your role below and follow the recommended reading path.

## 🎯 Quick Navigation by Role

### 👔 **Executive / Senior Stakeholder**

**Your Goal:** Understand business value, ROI, and strategic direction

**Start Here:** 1. 📊 [Executive Summary](01_EXECUTIVE_SUMMARY/executive_summary.md) *(10 minutes)* - Problem statement & solution - Implementation phases & timeline - Cost-benefit analysis - Key metrics & KPIs - Risks & recommendations

**That’s it!** You have everything you need to make a decision.

### 🎯 **Product Owner**

**Your Goal:** Understand requirements, scope, and business impact

**Recommended Reading Order:** 1. 📊 [Executive Summary](01_EXECUTIVE_SUMMARY/executive_summary.md) *(10 min)* - Business case and ROI

1. 📋 [Video Tracking Scenarios Guide](02_BUSINESS_ANALYSIS/VIDEO_TRACKING_SCENARIOS_GUIDE.md) *(30 min)*
   * All 10 tracking scenarios explained
   * Business insights for each scenario
   * Metric definitions
   * Use cases
2. 📖 <README.md> *(10 min)*
   * Solution overview
   * Features & capabilities

**Optional Deep Dive:** - 🔍 [Visual Guide - Closing Events](02_BUSINESS_ANALYSIS/VISUAL_GUIDE_CLOSING_EVENTS.md) - Why some watch time is lost

**Time Investment:** 50 minutes to full understanding

### 📊 **Business Analyst**

**Your Goal:** Define requirements, validate data, and ensure metrics align with business needs

**Recommended Reading Order:** 1. 📋 [Video Tracking Scenarios Guide](02_BUSINESS_ANALYSIS/VIDEO_TRACKING_SCENARIOS_GUIDE.md) *(45 min)* - **All 10 scenarios** with examples - Raw data → Transformed output - Metric calculations explained - Data quality considerations - Business use cases

1. 🔍 [Visual Guide - Closing Events](02_BUSINESS_ANALYSIS/VISUAL_GUIDE_CLOSING_EVENTS.md) *(20 min)*
   * Deep dive on event pairs
   * Why browser close matters
   * Max position vs watch time
2. 📊 [Executive Summary](01_EXECUTIVE_SUMMARY/executive_summary.md) *(10 min)*
   * Overall context and phases
3. 📖 [Quick Reference Guide](05_REFERENCE/quick_reference_guide.md) *(Reference)*
   * Complete metric definitions
   * Schema reference

**Your Deliverables:** - Requirements document (based on scenarios) - Acceptance criteria (based on expected outputs) - Test case validation (compare with scenario examples)

**Time Investment:** 1.5 hours for complete understanding

### 🏃 **Scrum Master**

**Your Goal:** Understand scope, plan sprints, identify dependencies

**Recommended Reading Order:** 1. 📊 [Executive Summary](01_EXECUTIVE_SUMMARY/executive_summary.md) *(15 min)* - See **Implementation Phases** section - 4 phases, each 2 weeks

1. 📋 [Video Tracking Scenarios Guide](02_BUSINESS_ANALYSIS/VIDEO_TRACKING_SCENARIOS_GUIDE.md) *(30 min)*
   * Understand complexity
   * Identify edge cases for sprint planning
2. 🚀 [Getting Started Guide](05_REFERENCE/GETTING_STARTED.md) *(10 min)*
   * Implementation steps
   * Dependencies

**Sprint Planning Help:** - **Phase 1 (Sprint 1-2):** Basic scenarios (1-2 from guide) - **Phase 2 (Sprint 3-4):** Edge cases (3, 4, 5) - **Phase 3 (Sprint 5-6):** Advanced scenarios (6-10) - **Phase 4 (Sprint 7-8):** Performance optimization

**Time Investment:** 1 hour

### 👨‍💻 **Developer / Data Engineer**

**Your Goal:** Implement the solution correctly and efficiently

**Recommended Reading Order:** 1. 📚 [**Documentation Summary**](DOCUMENTATION_SUMMARY.md) *(10 min)* ⭐ **START HERE** - Overview of all code documentation - Quick reference by topic - Finding specific information

1. 📖 [**Code Documentation Guide**](03_DEVELOPMENT/CODE_DOCUMENTATION_GUIDE.md) *(20 min)*
   * Documentation standards
   * How code is documented
   * Examples of good documentation
   * Checklist for code reviews
2. 🚀 [Getting Started Guide](05_REFERENCE/GETTING_STARTED.md) *(30 min)*
   * Step-by-step implementation
   * Prerequisites
   * Code setup
3. 💻 [Main Script: databricks\_video\_aggregation.py](03_DEVELOPMENT/databricks_video_aggregation.py)
   * **The implementation code**
   * **Extensively documented with examples**
   * Every method has algorithm explanations
   * Includes expected outcomes and calculations
   * Copy and customize
4. 📓 [Example Notebook: databricks\_example\_notebook.py](03_DEVELOPMENT/databricks_example_notebook.py)
   * Test data generation with detailed comments
   * Validation examples with expected results
   * Sample queries explained
5. 📖 [Quick Reference Guide](05_REFERENCE/quick_reference_guide.md) *(Reference)*
   * Complete technical reference
   * All functions explained
   * Troubleshooting

**Optional for Context:** - 📋 [Video Tracking Scenarios Guide](02_BUSINESS_ANALYSIS/VIDEO_TRACKING_SCENARIOS_GUIDE.md) - Understand what you’re building

**Key Documentation Features:** - ✅ Every file has comprehensive header documentation (100+ lines) - ✅ All methods include algorithm explanations with examples - ✅ Concrete calculations shown (e.g., “watchPercentage = 130/300 = 43.3%”) - ✅ Business context for every metric - ✅ Data quality rules explained - ✅ Debugging tips included - ✅ 90%+ of code has inline comments

**Time Investment:** 2-3 hours to understand fully, 2 hours to implement

### 🧪 **QA / Tester**

**Your Goal:** Validate all scenarios work correctly, ensure data quality

**Recommended Reading Order:** 1. 📚 [**Documentation Summary**](DOCUMENTATION_SUMMARY.md) *(10 min)* ⭐ **START HERE** - Overview of test documentation - Quick reference for finding test cases - Expected outcomes by scenario

1. 🧪 [**Complete Test Data Generator**](04_TESTING/test_data_generator_complete.py) *(30 min)* ⭐ **MAIN TEST SUITE**
   * **25 comprehensive test scenarios** (10 core + 15 edge cases)
   * Each scenario includes:
     + Description of what it tests
     + Timeline of events
     + Expected metrics with calculations
     + Why the test matters
   * Covers 90%+ of real-world cases
   * Runnable test data generator
2. 📋 [Video Tracking Scenarios Guide](02_BUSINESS_ANALYSIS/VIDEO_TRACKING_SCENARIOS_GUIDE.md) *(60 min)*
   * **Core 10 business scenarios**
   * Input data (raw events)
   * Expected output (metrics)
   * Business validation
3. 📓 [Example Notebook](03_DEVELOPMENT/databricks_example_notebook.py) *(30 min)*
   * Test data generation scripts with detailed comments
   * Validation queries with expected results
   * Step-by-step examples
4. 🔍 [Visual Guide - Closing Events](02_BUSINESS_ANALYSIS/VISUAL_GUIDE_CLOSING_EVENTS.md) *(20 min)*
   * Edge cases to test
   * Browser close scenarios

**Your Test Plan (25 Test Cases):**

CORE SCENARIOS (TC-001 to TC-010):  
Test Case 1: Perfect Viewing → Scenario 1  
Test Case 2: Pause & Resume → Scenario 2  
Test Case 3: Browser Close → Scenario 3  
Test Case 4: Skip Forward → Scenario 4  
Test Case 5: Rewind → Scenario 5  
Test Case 6: Multiple Sessions → Scenario 6  
Test Case 7: Multi-Video → Scenario 7  
Test Case 8: Early Abandonment → Scenario 8  
Test Case 9: Complex Navigation → Scenario 9  
Test Case 10: Gaming Detection → Scenario 10  
  
EDGE CASES (TC-011 to TC-025):  
Test Case 11: Duplicate Events  
Test Case 12: Out-of-Order Events  
Test Case 13: Null/Missing Values  
Test Case 14: Negative currentTime  
Test Case 15: Extremely Long Watch  
... (see test generator for complete list)

**Key Testing Features:** - ✅ 25 test scenarios documented with expected outcomes - ✅ Each test includes timeline and metric calculations - ✅ Runnable test data generator - ✅ Validation queries included - ✅ Covers data quality edge cases

**Validation Queries:** See test generator and example notebook.

**Time Investment:** 2 hours to understand, then test execution

## 📁 Repository Structure

VideoAnalytics/  
│  
├── INDEX.md (THIS FILE)  
│ └── Your navigation guide  
│  
├── README.md  
│ └── Project overview for everyone  
│  
├── DOCUMENTATION\_SUMMARY.md ⭐ NEW! COMPREHENSIVE DOCS GUIDE  
│ └── Overview of all code documentation  
│ - What's been enhanced  
│ - Documentation style  
│ - Quick start by role  
│ - Finding specific information  
│  
├── 01\_EXECUTIVE\_SUMMARY/  
│ └── executive\_summary.md  
│ └── For senior stakeholders & decision makers  
│  
├── 02\_BUSINESS\_ANALYSIS/  
│ ├── VIDEO\_TRACKING\_SCENARIOS\_GUIDE.md ⭐ MAIN GUIDE  
│ │ └── All scenarios explained (BA, PO, Testers)  
│ └── VISUAL\_GUIDE\_CLOSING\_EVENTS.md  
│ └── Deep dive on event pairs  
│  
├── 03\_DEVELOPMENT/  
│ ├── CODE\_DOCUMENTATION\_GUIDE.md ⭐ NEW! DOCUMENTATION STANDARDS  
│ │ └── How to document code  
│ │ - File/class/method templates  
│ │ - Examples of good documentation  
│ │ - Checklist for code reviews  
│ │  
│ ├── databricks\_video\_aggregation.py ⭐ MAIN CODE (EXTENSIVELY DOCUMENTED)  
│ │ └── The implementation with:  
│ │ - 120+ line header documentation  
│ │ - Algorithm explanations with examples  
│ │ - Expected outcomes & calculations  
│ │ - Business context for every metric  
│ │ - Debugging tips  
│ │  
│ └── databricks\_example\_notebook.py (EXTENSIVELY DOCUMENTED)  
│ └── Tutorial with:  
│ - Step-by-step examples  
│ - Detailed comments for each scenario  
│ - Expected results explained  
│  
├── 04\_TESTING/  
│ └── test\_data\_generator\_complete.py ⭐ COMPREHENSIVE TEST SUITE  
│ └── 25 test scenarios (90%+ coverage):  
│ - 10 core scenarios  
│ - 15 edge cases  
│ - Each with timeline & expected metrics  
│ - Runnable test data generator  
│  
└── 05\_REFERENCE/  
 ├── GETTING\_STARTED.md ⭐ SETUP GUIDE  
 │ └── Step-by-step implementation  
 ├── quick\_reference\_guide.md  
 │ └── Technical reference  
 └── QUICK\_REFERENCE\_CARD.md  
 └── Quick lookup

**🆕 New Documentation Files:** - **DOCUMENTATION\_SUMMARY.md**: Your starting point for understanding all documentation - **CODE\_DOCUMENTATION\_GUIDE.md**: Standards for maintaining clear documentation - **Enhanced code files**: All Python files now have extensive inline documentation

## 🚀 Typical Workflow

### **Phase 1: Discovery & Analysis** (Week 1)

**Who:** Product Owner, Business Analyst, Scrum Master

**Activities:** 1. PO reads Executive Summary & Scenarios Guide 2. BA reads Scenarios Guide & creates requirements doc 3. Scrum Master plans sprints based on phases 4. Team reviews requirements together

**Output:** - ✅ Requirements document - ✅ Acceptance criteria - ✅ Sprint plan

### **Phase 2: Development** (Week 2-7)

**Who:** Developers, Business Analyst (for questions)

**Activities:** 1. Devs read Getting Started Guide 2. Setup Databricks environment 3. Implement using provided scripts 4. Test with example notebook 5. BA validates scenarios 1-3 (Sprint 1) 6. Continue through all scenarios

**Output:** - ✅ Working aggregation script - ✅ Scheduled daily job - ✅ Aggregated data table

### **Phase 3: Testing & Validation** (Week 8)

**Who:** Testers, Business Analyst, Developers

**Activities:** 1. QA reads Scenarios Guide 2. Create test cases (1 per scenario) 3. Run validation queries 4. BA validates business logic 5. Fix any issues

**Output:** - ✅ Test report (10/10 scenarios pass) - ✅ Data quality validation - ✅ Sign-off from BA

### **Phase 4: Deployment & Reporting** (Week 9-10)

**Who:** Developers, Product Owner, Business Analyst

**Activities:** 1. Deploy to production 2. Connect to BI tool (Tableau/Power BI) 3. Create dashboards 4. PO reviews with stakeholders 5. Gather feedback

**Output:** - ✅ Production deployment - ✅ Dashboards live - ✅ Executive presentation

## 📊 Key Documents by Purpose

### **For Understanding Business Value:**

* [Executive Summary](01_EXECUTIVE_SUMMARY/executive_summary.md) - ROI, phases, timeline

### **For Understanding What to Build:**

* [Video Tracking Scenarios Guide](02_BUSINESS_ANALYSIS/VIDEO_TRACKING_SCENARIOS_GUIDE.md) - Complete requirements

### **For Implementation:**

* [Getting Started Guide](05_REFERENCE/GETTING_STARTED.md) - Setup steps
* [databricks\_video\_aggregation.py](03_DEVELOPMENT/databricks_video_aggregation.py) - The code

### **For Testing:**

* [Video Tracking Scenarios Guide](02_BUSINESS_ANALYSIS/VIDEO_TRACKING_SCENARIOS_GUIDE.md) - Test cases
* [databricks\_example\_notebook.py](03_DEVELOPMENT/databricks_example_notebook.py) - Test data

### **For Reference:**

* [Quick Reference Guide](05_REFERENCE/quick_reference_guide.md) - Complete technical docs

## 🎯 Success Criteria by Role

### **Executives:**

* ✅ Understand business value (10 min)
* ✅ Approve budget/timeline
* ✅ Receive quarterly reports

### **Product Owner:**

* ✅ Understand all scenarios (50 min)
* ✅ Define acceptance criteria
* ✅ Validate business logic
* ✅ Present insights to stakeholders

### **Business Analyst:**

* ✅ Master all 10 scenarios (90 min)
* ✅ Create requirements document
* ✅ Validate data transformations
* ✅ Define KPIs

### **Scrum Master:**

* ✅ Understand implementation phases (60 min)
* ✅ Plan 4-8 sprints
* ✅ Identify dependencies
* ✅ Track progress

### **Developers:**

* ✅ Setup environment (30 min)
* ✅ Implement code (2 hours)
* ✅ Handle all scenarios
* ✅ Pass all tests

### **Testers:**

* ✅ Understand 10 scenarios (90 min)
* ✅ Create 10+ test cases
* ✅ Validate all outputs
* ✅ Sign-off on quality

## 🎓 Recommended Reading Time

| **Role** | **Essential Reading** | **Time** | **Optional Reading** | **Total** |
| --- | --- | --- | --- | --- |
| **Executive** | Executive Summary | 10 min | - | 10 min |
| **Product Owner** | 3 docs | 50 min | Visual Guide | 70 min |
| **Business Analyst** | 4 docs | 90 min | Dev code | 120 min |
| **Scrum Master** | 3 docs | 60 min | - | 60 min |
| **Developer** | 4 docs | 120 min | Scenarios | 180 min |
| **Tester** | 3 docs | 110 min | - | 110 min |

## 💡 Pro Tips

### **For Product Owners:**

* Start with Executive Summary for context
* Then dive deep into Scenarios Guide - this is your requirements doc!
* Use scenarios to write user stories

### **For Business Analysts:**

* The Scenarios Guide is your main reference
* Each scenario shows input → output transformation
* Use these as validation examples
* Compare your data with expected outputs

### **For Developers:**

* Don’t skip Getting Started - it has critical setup steps
* The code is well-commented - read it!
* Run example notebook first before using real data
* Reference guide has troubleshooting section

### **For Testers:**

* Each scenario in the guide = 1 test case
* Input data is provided (raw events)
* Expected output is shown (metrics)
* Validation queries are included

### **For Scrum Masters:**

* Phases in Executive Summary = Sprint planning guide
* Scenarios increase in complexity - plan accordingly
* Phase 1-2 are MVP, Phase 3-4 are nice-to-have

## 🆘 Need Help?

### **“I don’t know where to start”**

→ Find your role at the top of this document

### **“I need to understand the business case”**

→ [Executive Summary](01_EXECUTIVE_SUMMARY/executive_summary.md)

### **“I need to understand what we’re building”**

→ [Video Tracking Scenarios Guide](02_BUSINESS_ANALYSIS/VIDEO_TRACKING_SCENARIOS_GUIDE.md)

### **“I need to implement this”**

→ [Getting Started Guide](05_REFERENCE/GETTING_STARTED.md)

### **“I need to test this”**

→ [Video Tracking Scenarios Guide](02_BUSINESS_ANALYSIS/VIDEO_TRACKING_SCENARIOS_GUIDE.md) + [Example Notebook](03_DEVELOPMENT/databricks_example_notebook.py)

### **“I need technical details”**

→ [Quick Reference Guide](05_REFERENCE/quick_reference_guide.md)

## 📞 Document Ownership

| **Document** | **Primary Audience** | **Maintained By** |
| --- | --- | --- |
| Executive Summary | Executives, PO | Product Owner |
| Scenarios Guide | BA, PO, Testers | Business Analyst |
| Code Files | Developers | Development Team |
| Getting Started | Developers | Technical Lead |
| Visual Guide | BA, Testers | Business Analyst |
| Quick Reference | Developers | Technical Lead |

## 🎯 Next Steps

**Choose your role above and start reading!**

Most roles can be productive in **under 2 hours** of reading.

## ✅ Checklist by Role

### **Executive:**

* Read Executive Summary
* Approve timeline & budget
* Schedule follow-up for results

### **Product Owner:**

* Read Executive Summary
* Read Scenarios Guide
* Define acceptance criteria
* Review with stakeholders

### **Business Analyst:**

* Read Scenarios Guide (all 10 scenarios)
* Read Visual Guide
* Create requirements document
* Define validation criteria

### **Scrum Master:**

* Read Executive Summary (phases)
* Read Scenarios Guide (complexity)
* Create sprint plan (4-8 sprints)
* Identify dependencies

### **Developer:**

* Read Getting Started Guide
* Setup Databricks environment
* Run example notebook
* Implement solution
* Pass all test scenarios

### **Tester:**

* Read Scenarios Guide (all 10)
* Create test cases (1 per scenario)
* Run validation queries
* Document test results
* Sign-off on quality

## 🌟 Why This Structure?

**Problem:** Too many files, unclear who should read what.

**Solution:** - ✅ Clear folder structure (by phase) - ✅ Role-based navigation (this file!) - ✅ Estimated reading times - ✅ Recommended reading order - ✅ Clear deliverables per role

**Result:** Everyone knows exactly what to read and why.

**Document Version:** 1.0 **Last Updated:** 2025-10-06 **Maintained By:** Project Lead

*Made with 📚 for better team collaboration*