 # ECON485 Project 3 - Stage 2 Presentation

Tourism Revenue & Visitor Trends Database

****Team 3****

****Week 9, Fall 2025****

****Stage 2: Design & Prototype****

📋 Presentation Agenda

1. ****Project Recap**** (2 min)
2. ****Schema Design & Normalization**** (3 min)
3. ****Data Model Demonstration**** (4 min)
4. ****Business Query Results**** (5 min)
5. ****Challenges & AI Learning**** (3 min)
6. ****Next Steps**** (2 min)
7. ****Q&A**** (1 min)

****Total: 20 minutes****

1️⃣ PROJECT RECAP

Tourism Analytics System

****Objective:**** Enable data-driven tourism policy and business decisions

Scope:

- Track visitor demographics and spending patterns

- Analyze hotel occupancy and seasonal trends

- Measure regional economic impact

- Support pricing and marketing strategies

Stakeholders:

- Tourism boards (policy planning)

- Hotel chains (revenue management)

- Regional governments (infrastructure investment)

2️⃣ SCHEMA DESIGN & NORMALIZATION

Entity-Relationship Overview

7 Core Tables:

1. **Regions** - Geographic destinations

2. **Hotels** - Accommodation facilities

3. **Visitors** - Tourist demographics

4. **RoomTypes** - Accommodation categories

5. **Seasons** - Temporal classifications

6. **Bookings** - Reservations (transactional)

7. **Expenditures** - Spending records

Key Relationships

- Regions (1) → (N) Hotels

- Hotels (1) → (N) Bookings

- Visitors (1) → (N) Bookings

- Visitors (1) → (N) Expenditures

2️⃣ **NORMALIZATION STATUS**

3NF Compliance

✅ ****Fully Normalized:**** 5 tables (Regions, Visitors, RoomTypes, Seasons, Expenditures)

⚠️ ****Intentional Denormalization:**** 2 tables

1. ****Hotels:**** Address components (City, ZipCode)

  - Justified: Convenience, minimal update risk

2. ****Bookings:**** TotalCost (derived field)

  - Justified: Performance, historical accuracy, immutable bookings

Design Decisions

- All denormalizations documented with business rationale

- Foreign key constraints enforce referential integrity

- Check constraints validate business rules

- Indexes optimize query performance

3️⃣ **DATA MODEL DEMONSTRATION**

Sample Data Overview

Entity	Record Count	Date Range
Regions	10	—
Hotels	15	—
Visitors	16	—
RoomTypes	5	—
Seasons	8	2024-2025
Bookings	16	2024-07 to 2025-08
Expenditures	17	2024-07 to 2024-10

Data Characteristics

- **Geographic Coverage:** 10 Turkish regions (urban, coastal, rural, mountain)
- **Visitor Mix:** 50% domestic, 50% international (8 countries represented)
- **Seasonality:** Full year coverage with peak/shoulder/off-season distribution
- **Hotel Diversity:** 3.0 to 5.0 star ratings, 45-400 rooms

3️⃣ CRUD OPERATIONS SHOWCASE

CREATE (Insert)

✅ 16 visitors inserted with demographics

✅ 16 bookings across 2-year period

✅ 17 expenditure records across categories

READ (Select)

✅ Multi-table JOINs verified

✅ Aggregate queries functional

✅ Filtering and sorting tested

UPDATE

✅ Booking status updates (past → completed)

✅ Visitor information corrections

✅ Hotel capacity adjustments

DELETE

✅ Foreign key constraints enforced

✅ Orphaned record prevention confirmed

✅ Transaction rollback capability validated

🛂 BUSINESS QUERY RESULTS

Query 1: Regional Spending Analysis

Findings:

- **Istanbul:** Highest per-visitor spending (505 TL)

- **Cappadocia:** Premium activity spending (255 TL)

- **Antalya:** High volume, moderate per-visitor (377 TL)

Business Insight:

Urban and experience-based regions generate highest returns

Economic Implication:

Infrastructure investment should prioritize high-yield destinations

---

Query 2: Seasonal Hotel Performance

Findings:

- **Peak Season (Jun-Aug, Dec):** 60% of bookings, highest revenue
- **Shoulder Season (Apr-May, Sep-Oct):** 30% bookings, moderate pricing
- **Off-Season (Feb-Mar, Nov):** 10% bookings, significant discounts needed

Business Insight:

Severe seasonality creates revenue volatility

Hotel Strategy:

- Dynamic pricing essential (up to 3x peak vs off-season)
- Staff flex scheduling required
- Off-season promotions critical for cash flow

Query 3: Visitor Segmentation

Findings:

Visitor Type	Avg Spending/Visitor	Avg Stay Duration

| International | 2,266 TL | 6.5 days |

| Domestic | 892 TL | 4.2 days |

Key Differences:

- International visitors: 2.5x higher spending

- Longer stays correlate with higher total spending

- Room type preferences: International prefers suites/deluxe

Marketing Implication:

Target high-spending international markets (UK, Germany, Gulf states)

4️⃣ ECONOMIC INSIGHTS

Tourism Impact Analysis

Direct Effects:

- Accommodation revenue: 30,000+ TL (sample period)

- Non-accommodation spending: 2,500+ TL

- Total visitor spending: 32,500+ TL (16 visitors)

Extrapolated Annual Impact (1,000 visitors):

- Projected annual revenue: ~2 million TL per region

- Employment: ~15-20 direct jobs per hotel

- Multiplier effect: 1.8x (indirect/induced impact)

Regional GDP Contribution:

Tourism represents 8-15% of regional GDP in coastal areas

5️⃣ CHALLENGES & AI LEARNING

Technical Challenges

Challenge 1: Seasonal Classification

- **Issue:** How to link Seasons to Bookings (no direct FK)
- **Solution:** Derived relationship using CASE statements in queries
- **AI Tool Used:** ChatGPT suggested CTE approach
- **Learning:** Database design isn't always about direct relationships

Challenge 2: Occupancy Calculation

- **Issue:** True occupancy requires daily room inventory
- **Solution:** Simplified metric (bookings/total rooms)
- **AI Tool Used:** Perplexity AI researched hotel KPI standards
- **Limitation:** Our metric approximates but doesn't capture multi-night stays

Challenge 3: Sample Data Realism

- **Issue:** AI-generated data had unrealistic patterns
- **Solution:** Manual curation + Python script refinements
- **AI Tool Used:** ChatGPT Python generator
- **Correction:** Fixed visitor booking coverage, adjusted seasonal distribution

---

AI Tool Usage Summary

| Tool | Purpose | Success Rate | Key Learning |

|------|---------|--------------|--------------|

| **ChatGPT** | Schema design, query generation | 70% | Requires significant refinement |

| **Perplexity AI** | Industry research | 90% | Excellent for context/validation |

| **dbdiagram.io** | ER visualization | 95% | Intuitive, reliable |

| **GitHub Copilot** | SQL boilerplate | 80% | Good for CRUD, weak on business logic |

Overall Assessment:

- AI accelerates initial drafts (50% time savings)

- Human review essential for accuracy (71% of AI outputs required corrections)

- Domain expertise (economics + tourism) critical for meaningful analysis

---

Fail-Forward Examples

Failure #1: AI Suggested Wrong Normalization

- ChatGPT initially suggested keeping HotelName in Bookings

- **Error:** Violates 3NF (transitive dependency)

- **Correction:** Team caught during normalization review

- **Learning:** Always validate AI against database theory

Failure #2: Incomplete Sample Data

- Python script left some visitors without bookings

- **Error:** Unrealistic (all visitors should book hotels)

- **Correction:** Modified script to ensure 1:N relationship

- **Learning:** Test data distributions, not just syntax

Failure #3: Query Performance Assumption

- AI suggested removing TotalCost for "purity"

- **Error:** Would severely impact query performance

- **Correction:** Team justified intentional denormalization

- **Learning:** Theory vs practice trade-offs require business judgment

---

6️⃣ NEXT STEPS (Week 10-14)

Stage 3: Final Presentation Phase

Week 10: Advanced Queries (5+ complex JOINs)

- Revenue trends over time

- Occupancy forecasting

- Visitor segmentation deep-dive

- Regional competitiveness analysis

Week 11: Views & Indexes

- Create materialized views for dashboards

- Optimize slow queries with indexes

- Performance benchmarking (EXPLAIN)

Week 12-13: Integration & Visualization

- AI-assisted data visualization (recommendations)

- Economic narrative development

- GitHub repository finalization

- 150-word AI reflection paragraph

****Week 14:**** Final Demo

- Live database demonstration

- Economic policy recommendations

- Q&A with instructor

6️⃣ **EXPECTED DELIVERABLES**

Stage 3 Requirements

✅ ****Functional Integration:**** Working prototype with KPI calculations

✅ ****Visualization:**** AI-assisted charts showing seasonal trends

✅ ****Economic Interpretation:**** Policy implications document

✅ ****Complete Repository:**** All code, data, docs on GitHub

✅ ****AI Reflection:**** 150-word learning summary

Evaluation Criteria (45% of project grade)

- Functionality: Does it calculate correct KPIs?

- Insight: Clear economic interpretation?

- Presentation: Professional slides and demo?

- Documentation: Complete GitHub repository?

7️⃣ TEAM CONTRIBUTIONS

Individual Roles & Responsibilities

Mert Ak□n - Project Lead:

- Schema design and normalization
- GitHub repository management
- Team coordination and deadlines

[Member 2] - Data Analyst:

- SQL query development and optimization
- Sample data generation (Python scripts)
- Performance testing

[Member 3] - Economic Analyst:

- Business logic and KPI definitions
- Economic insights and interpretation
- Industry research

[Member 4] - Documentation Lead:

- AI interaction logging
- Meeting notes and progress reports
- Presentation slide creation

Team Collaboration: All members participated in design discussions, code reviews, and deliverable preparation

---

📊 PROGRESS SUMMARY

Stage 1 (Week 3-4): ✅ Completed

- Project definition document
- Initial ER diagram
- GitHub repository setup

Stage 2 (Week 5-9): ✅ Completed

- Normalized schema (3NF)
- SQL implementation with CRUD operations
- 3 business analysis queries
- Sample data (85+ records)
- This presentation

Stage 3 (Week 10-14): 🔄 In Progress

- Advanced queries
- Views and indexes
- Final integration
- Economic recommendations

💡 KEY TAKEAWAYS

Technical Learnings

1. **Normalization is a judgment call:** Theory provides guidelines, business needs determine exceptions

2. **Query performance matters:** Denormalization justified when performance critical

3. **Data quality > quantity:** 100 realistic records beat 10,000 unrealistic ones

AI Learnings

4. **AI accelerates, doesn't replace:** 50% time savings, but human oversight essential

5. **Domain expertise crucial:** Economics + tourism knowledge caught AI errors

6. **Documentation valuable:** Logging AI interactions revealed patterns

Business Learnings

7. **Seasonality dominates:** Turkish tourism highly seasonal → revenue volatility

8. **International premium:** Foreign tourists drive disproportionate revenue

9. **Data-driven decisions:** Quantitative analysis supports policy recommendations

---

❓ Q&A

Questions?

Contact:

- GitHub: [Repository Link]

- Slack: ECON485-Team3-Tourism

- Email: [team lead email]

Thank you!

APPENDIX: Sample Query Output

Query 1 Results

```
RegionName   | TotalVisitors | TotalRevenue | AvgSpendingPerVisitor
-------------|---------------|--------------|----------------------
Istanbul     | 3             | 1515.00      | 505.00
Cappadocia   | 1             | 255.00       | 255.00
Antalya      | 2             | 755.00       | 377.50
```

Query 2 Results

```
HotelName           | Season    | TotalBookings | Revenue
--------------------|-----------|---------------|----------
Antalya Beach Resort  | Peak      | 2             | 6290.00
Grand Istanbul Palace | Peak      | 1             | 1750.00
```

Query 3 Results

```
Country       | VisitorType   | SpendingPerVisitor
--------------|---------------|-------------------
United Kingdom | International | 3390.00
Germany       | International | 2240.00
```

Turkey | Domestic | 892.00

```

---

**\*\*END OF PRESENTATION\*\***

*\*Stage 2: Design & Prototype - Completed Successfully\**

*\*Next Milestone: Week 10 - Advanced Queries\**
```