Data and Reporting Engineers

Training track

Data engineers, analytics and BI engineers, regulatory reporting developers, finance data analysts

You will be able to build a regulator report that ties back to the ledger, arrives on time and can be proven correct a year later.

- Recommended path
- Optional, worth it
- Alternative route

DURATION	SLIDES	FORMAT	PREREQUISITES	OUTCOME	REGIONS
16.9 h of core lessons + 120 min of extension…	237 in lessons + 16 extension + 12-18 for the…	Seven to eight remote sessions (never more than…	Two years of data work; strong SQL, working Python.	A working evidence export and a reporting pipeline…	One region, fixed: the regulator feed pair 14.4…

STAGE 1 · Session 1 — Foundations and data flows

- 1.1 Regulated system and the method
- 1.2 Bet lifecycle and data flows

STAGE 2 · Sessions 1-3 — The ledger as source of truth

- 2.1 Wallet, double entry, accounts
- 2.3 Payments, treasury and reconciliation
- 2.2 Lab: post, park and reverse
- 2.4 Design: PSP mismatch

STAGE 3 · Sessions 3-4 — Regulated data sets

- 3.1 RTP and game math data
- 3.3 Behavior monitoring data
- 3.2 Evidence store and AML data

STAGE 4 · Sessions 4-5 — Data placement

- 4.1 Classes, hub and spoke, transfers
- 4.2 Design: three-market placement

STAGE 5 · Sessions 5-7 — Reporting and evidence

- 5.1 Evidence pipeline and formats
- 5.3 Frameworks, calendar, gates
- 5.5 Deletion vs retention
- 5.2 Lab: evidence export
- 5.4 Design: batch pipeline
- 5.6 Audit logs and SIEM

STAGE 6 · Sessions 7-8 — Region elective (pick one)

- 6.1 Region: Latin America
- 6.2 Region: US states

STAGE 7 · Session 8 — Track extension lessons

- 7.1 Data contract, regulator feed
- 7.3 Reconciliation lab: GGR to ledger
- 7.2 Tie report totals to ledger

Data and Reporting Engineers: training track in detail

Data engineers, analytics and BI engineers, regulatory reporting developers, finance data analysts · as of 2026-11

Duration	16.9 h of core lessons + 120 min of extension lessons + one region (110-125 min) + 35-40 min of session openings: 21.3-21.7 h in all
Slides	237 in lessons + 16 extension + 12-18 for the region: 265-271 in all
Format	Seven to eight remote sessions (never more than 3.5 hours of contact each), or four days on site. Three labs (post and reverse, evidence export, reconciliation) and whiteboard designs. About 6 hours of post-work for the closing task.
Prerequisites	Two years of data work; strong SQL, working Python.
Outcome	A working evidence export and a reporting pipeline design that reconciles to the ledger.
Regions	One region, fixed: the regulator feed pair 14.4 and 14.5. A US cohort swaps in 12.3a, 12.3b and 12.5.
Assessment	Closing quiz (20 questions). Labs under the 75 percent rule. Closing task: a data contract and one report reconciled to the ledger with the differences explained.
Certificate	Attendance record only. Hours count toward the full program.

1. Foundations and data flows Session 1

1.1 Regulated system and the method RECOMMENDED

The business and the five-step method. Data engineers own much of the evidence step. Lessons 1.1 and 1.3: 75 min, 20 slides.

In the book: Ch. 01 The Online Casino Ecosystem, Ch. 05 Differences Between Betting Sites and Online Casinos, Ch. 07 Casino Implementation Planning and Timeline

1.2 Bet lifecycle and data flows RECOMMENDED

Where each event of a bet is produced and which store is the source of truth for it. Lesson 2.2: 35 min, 12 slides.

In the book: Ch. 10 Complete Platform Architecture, Ch. 28a Distributed Systems Deep Dive

2. The ledger as source of truth Sessions 1-3

2.1 Wallet, double entry, accounts RECOMMENDED

The wallet rules a report may rely on, then the double-entry ledger and its chart of accounts. Every financial report starts here. Lessons 3.1 and 3.2: 80 min, 24 slides.

In the book: Ch. 36 Financial Operations, Ch. 10 Complete Platform Architecture, Ch. 46 Building a Brazilian Betting Platform: Architecture, Compliance, and Implementation, Ch. 36b The Financial Truth Layer -- Ledger, Treasury, and Reconciliation

2.2 Lab: post, park and reverse RECOMMENDED

Post, park and reverse entries and prove the trial balance, in SQL. Lesson 3.3: 65 min, 4 slides.

In the book: Ch. 36b The Financial Truth Layer -- Ledger, Treasury, and Reconciliation, Ch. 36 Financial Operations

2.3 Payments, treasury and reconciliation RECOMMENDED

Payment layers and the files a provider sends, then float, segregation of player funds and the daily reconciliation. Lessons 3.4 and 3.5: 80 min, 22 slides.

In the book: Ch. 36 Financial Operations, Ch. 46 Building a Brazilian Betting Platform: Architecture, Compliance, and Implementation, Ch. 33d Disputes, Chargebacks, and Player Support, Ch. 36b The Financial Truth Layer -- Ledger, Treasury, and Reconciliation

2.4 Design: PSP mismatch RECOMMENDED

Design the reconciliation for a mismatch between the ledger and a payment provider file. Lesson 3.6: 50 min, 5 slides.

In the book: Ch. 36b The Financial Truth Layer -- Ledger, Treasury, and Reconciliation

3. Regulated data sets Sessions 3-4

3.1 RTP and game math data RECOMMENDED

RTP and volatility, what gets certified and what has to be monitored in production data. Lesson 4.2: 30 min, 10 slides.

In the book: Ch. 15 Casino Mathematics and Game Economy, Ch. 34b Data Governance for iGaming Platforms

3.2 Evidence store and AML data RECOMMENDED

The KYC evidence store and the data that transaction monitoring needs. Lessons 5.2 and 5.4: 65 min, 20 slides.

In the book: Ch. 24d KYC Evidence Lifecycle, Ch. 20 Hardware Security Module Infrastructure, Ch. 20b OpenBao Operations: Secret Engines, Dynamic Credentials and Disaster Recovery, Ch. 02 Regulation and Compliance Landscape, Ch. 34b Data Governance for iGaming Platforms, Ch. 12 Real-Time Cash Flow Management for Online Casinos

3.3 Behavior monitoring data RECOMMENDED

Markers of harm as features and events, with the audit trail of each intervention. Lesson 6.4: 35 min, 11 slides.

In the book: Ch. 26 Responsible Gaming and Player Protection Systems

4. Data placement Sessions 4-5

4.1 Classes, hub and spoke, transfers RECOMMENDED

Data classification and residency rules, hub and spoke topology, then the gateway that decides what a central warehouse may receive. Lessons 8.1, 8.2 and 8.3: 125 min, 35 slides.

In the book: Ch. 27 Data Residency and Backup/Recovery, Ch. 10 Complete Platform Architecture, Ch. 27c Migrating a Single-Jurisdiction Casino Platform to Hub & Spoke, Ch. 27b The Jurisdiction Transfer Gateway and Cookie Consent, Ch. 34b Data Governance for iGaming Platforms

4.2 Design: three-market placement RECOMMENDED

Place every data class for a three-market operator. Data engineers argue for what analytics really needs. Lesson 8.7: 70 min, 5 slides.

In the book: Ch. 27 Data Residency and Backup/Recovery, Ch. 27c Migrating a Single-Jurisdiction Casino Platform to Hub & Spoke, Ch. 10 Complete Platform Architecture

5. Reporting and evidence Sessions 5-7

5.1 Evidence pipeline and formats RECOMMENDED

The evidence pipeline, then report formats and deadlines by regulator. Lessons 9.1 and 9.2: 75 min, 24 slides.

In the book: Ch. 25b Regulatory Reporting and Evidence Export, Ch. 06 Licensing Guide

5.2 Lab: evidence export RECOMMENDED

Build an export that writes an audit entry for every request. Lesson 9.3: 65 min, 4 slides.

In the book: Ch. 25b Regulatory Reporting and Evidence Export, Ch. 24d KYC Evidence Lifecycle

5.3 Frameworks, calendar, gates RECOMMENDED

Which audits look at data controls, and the calendar of report and audit dates. Lessons 9.4 and 9.5: 65 min, 22 slides.

In the book: Ch. 25 GLI-GSF Compliance Framework: Online Gaming Information Security, Ch. 24 Security and Compliance, Ch. 25b Regulatory Reporting and Evidence Export

5.4 Design: batch pipeline RECOMMENDED

A batch reporting pipeline with durable staging, signing and retries. Lesson 9.6: 50 min, 5 slides.

In the book: Ch. 46 Building a Brazilian Betting Platform: Architecture, Compliance, and Implementation, Ch. 25b Regulatory Reporting and Evidence Export

5.5 Deletion vs retention RECOMMENDED

A deletion request against a retention duty. Lesson 9.7: 20 min, 4 slides.

In the book: Ch. 34b Data Governance for iGaming Platforms, Ch. 25 GLI-GSF Compliance Framework: Online Gaming Information Security, Ch. 27 Data Residency and Backup/Recovery

5.6 Audit logs and SIEM RECOMMENDED

Audit logs as a data set with its own schema, retention and access rules. Lesson 10.3: 30 min, 10 slides.

In the book: Ch. 24b Wazuh SIEM for iGaming Compliance, Ch. 24c AWS SIEM Implementation for iGaming Compliance, Ch. 10 Complete Platform Architecture, Ch. 24f Network IDS/IPS: Suricata and AWS Network Firewall for iGaming

6. Region elective (pick one) Sessions 7-8

6.1 Region: Latin America ALTERNATIVE

Brazil, Colombia, Mexico and Buenos Aires. Brazil is taught as a design problem behind a legal status gate, because MP 1.394/2026 banned fixed-odds betting in September 2026. This track takes the Brazil regulator feed pair: signed batches to SIGAP and the lab that builds the pipeline. Brazil is taught as a design case behind the legal status gate. A US cohort takes the Pennsylvania and Michigan reporting lessons and the per-state spoke lab instead. Lessons 14.4 and 14.5: 110 min, 12 slides.

In the book: Ch. 46 Building a Brazilian Betting Platform: Architecture, Compliance, and Implementation, Ch. 25b Regulatory Reporting and Evidence Export, Ch. 26b Self-Exclusion Registries: From API to Audit Trail

6.2 Region: US states ALTERNATIVE

New Jersey, Pennsylvania and Michigan: a land-based partner, servers inside the state and a location check on every session. This track takes the Brazil regulator feed pair: signed batches to SIGAP and the lab that builds the pipeline. Brazil is taught as a design case behind the legal status gate. A US cohort takes the Pennsylvania and Michigan reporting lessons and the per-state spoke lab instead. Lessons 12.3a, 12.3b and 12.5: 125 min, 18 slides.

In the book: Ch. 06 Licensing Guide, Ch. 29a On-Premises Infrastructure for US-Regulated iGaming, Ch. 26b Self-Exclusion Registries: From API to Audit Trail, Ch. 24e Geofencing and Location Verification, Ch. 10 Complete Platform Architecture, Ch. 27 Data Residency and Backup/Recovery

7. Track extension lessons Session 8

7.1 Data contract, regulator feed OPTIONAL

Teams write a data contract for one regulator feed: schema, source tables, lineage, how late or corrected data is handled, who signs off. It follows the batch pipeline design. Not built yet; estimate 40 min, 4 slides (whiteboard design).

In the book: Ch. 25b Regulatory Reporting and Evidence Export, Ch. 34b Data Governance for iGaming Platforms, Ch. 34 Data and Analytics

7.2 Tie report totals to ledger OPTIONAL

A worked example that ties a monthly GGR report to ledger accounts, line by line, and lists the usual causes of a difference: time zones, voids, bonus money, currency. It is the brief for the reconciliation lab 3.E2. Not built yet; estimate 20 min, 8 slides (lecture).

In the book: Ch. 36b The Financial Truth Layer -- Ledger, Treasury, and Reconciliation, Ch. 25b Regulatory Reporting and Evidence Export

7.3 Reconciliation lab: GGR to ledger OPTIONAL

Tie a GGR report to the ledger in SQL: find the three planted differences, explain each one and write the correcting entry. The track promised reconciliation work and had only a lecture. Lab environment L22-reconciliation. Not built yet; estimate 60 min, 4 slides (lab).

In the book: Ch. 36b The Financial Truth Layer -- Ledger, Treasury, and Reconciliation, Ch. 25b Regulatory Reporting and Evidence Export

Common pitfalls

- Building reports from the analytics warehouse when the ledger is the source of truth.
- Copying player-level data into a central lake before checking residency.
- Treating a regulator feed like an internal dashboard: no schema version, no replay, no proof of delivery.
- Fixing a wrong number by overwriting it, which destroys the record of what was sent.
- Ignoring time zones and cut-off times, a frequent cause of a report that does not tie.

Sources

1. AGCO, Registrar's Standards for Internet Gaming (full text)
   https://agco.ca/en/book/export/html/245361
2. Brazil, Medida Provisória 1.394 of 25 September 2026 (Planalto)
   https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2026/Mpv/mpv1394.htm

3. Brazil, SPA/MF regulatory agenda 2026-2027
   https://www.gov.br/fazenda/pt-br/composicao/orgaos/secretaria-de-premios-e-apostas/agenda-regulatoria/2026
4. UK Gambling Commission, testing strategy for remote gambling and software technical standards, summary
   https://www.gamblingcommission.gov.uk/strategy/testing-strategy-for-compliance-with-remote-gambling-and-software-technical/1-summary
5. NIST SP 800-50 Rev. 1, Building a Cybersecurity and Privacy Learning Program
   https://csrc.nist.gov/pubs/sp/800/50/r1/final

Companion to *The Backend of Luck* and its jurisdiction roadmaps. Hours and slide counts come from the lesson tables in the training plan.