

OLIVER WYMAN

A MARSH BUSINESS

FAQ — Data Science Internship

Oliver Wyman — Internship

1. What is the main focus of this internship position?

The focus of this internship is to support **data analysis and insight generation** that informs **product, business, and operational decisions**. The work typically involves organizing datasets, defining metrics, performing exploratory analysis, and communicating findings clearly.

2. What does a Data Science intern do on a daily basis?

Daily work varies by team/project, but commonly includes:

- Exploring and understanding datasets (structure, meaning, gaps, limitations)
- Cleaning, organizing, and transforming data into analysis-ready form
- Defining and calculating metrics/KPIs with clear business context
- Performing exploratory data analysis (EDA) and identifying trends/drivers
- Building lightweight automations for recurring analyses in Python/SQL
- Creating reports, dashboards inputs, and short written summaries of insights
- Documenting logic and definitions so results are reproducible and consistent

3. Is this role more technical or more analytical?

The role is primarily **analytical**, with technical work as an enabler.

- **Analytical:** interpret data, form hypotheses, identify drivers, assess trade-offs
- **Technical:** use Python/SQL to query data, compute metrics, automate analysis, validate results

You won't be expected to do traditional software engineering, but you should be comfortable working hands-on with data.

4. Is this a programming role or a business/product role?

It's an **applied analytics role**. Programming (mainly **Python and SQL**) is used to explore data and build reliable metrics/analyses. The goal is to support decisions with evidence, not to build complex production systems.

5. Do I need prior experience with data to apply?

No prior professional experience is required. What matters most is:

- Basic familiarity with data analysis and descriptive statistics
- Ability to reason clearly from data
- Basic comfort with Python and SQL
- Curiosity and willingness to learn quickly

Projects and feedback loops are designed to help you grow in practice.

6. What level of Python and SQL is expected?

Basic to intermediate is expected.

Python

- Reading and cleaning datasets (pandas)
- Grouping/aggregation, simple transformations
- Creating summary tables and basic charts
- Writing clean, structured notebooks/scripts

SQL

- SELECT / JOIN / GROUP BY / CASE WHEN
- Basic filters, aggregations, sanity checks
- Window functions are a plus

If you can take a dataset from raw → clean → metrics → insights, you're well positioned.

7. Will I work with real client/company data? How is confidentiality handled?

Yes, interns typically work with **real project data** under supervision.

Confidentiality is critical in consulting. In practice, this means:

- Restricted access on a need-to-know basis
- Work done in approved environments (secure workspaces)
- Careful handling of extracts and outputs
- Sharing only through agreed internal channels

You'll receive clear guidance on policies and best practices.

8. What tools and environments will I use (e.g., notebooks, cloud, BI, Git)?

Tools vary by project, but commonly include:

- **Python notebooks** (for EDA, metrics, quick analysis)
- **SQL** (querying and shaping data)
- Cloud/Data platforms depending on the engagement
- **BI/Dashboard tools** (to visualize and share metrics, depending on the team)
- Slides and written summaries for communicating results
- Version control (e.g., Git) may appear depending on the setup

The key is being effective with Python/SQL and communicating findings clearly.

9. How are projects staffed at Oliver Wyman (one project vs multiple, team size, interaction with stakeholders)?

Staffing depends on project needs, but typically:

- You're assigned to a **specific team/project** with defined priorities
- You work closely with **more experienced professionals** who review and guide your work
- You often interact with consultants and sometimes with product/business stakeholders (through the team) to clarify questions, align assumptions, and interpret results

You'll learn how to deliver analysis in a consulting environment: structured, high-quality, and decision-oriented.

10. How are results typically presented (reports, dashboards, presentations)?

Most results are shared through:

- Concise **analytical write-ups** (what we found, why it matters, what to do next)
- Tables and charts (clear and decision-relevant)
- **Dashboards** or metric reporting (when recurring monitoring is needed)
- **Presentations** summarizing insights for stakeholders

Clarity, context, and decision impact are emphasized.

11. Is the work more exploratory or based on repetitive tasks?

Predominantly **exploratory**. You'll be exposed to open-ended questions like:

- "What's driving changes in this metric?"
- "Where do users drop off in the funnel?"
- "Which segments behave differently and why?"

There can also be recurring tasks (metric refreshes, QA checks), but interns often help automate and standardize these to reduce manual effort.

12. What type of profile stands out for this role?

Candidates who stand out usually demonstrate:

- Curiosity and strong analytical reasoning
- Organization and attention to detail (especially with definitions and data quality)
- Comfort forming hypotheses and iterating
- Clear communication (explaining assumptions, logic, and implications)
- Responsibility with deliverables and timelines

13. What type of data will I work with (structured, semi-structured, events/logs)?

You will primarily work with **structured and semi-structured data**, such as:

- Relational tables (operational data, product data, performance indicators)
- Time-series metrics and KPI tables
- Semi-structured event/log-like data depending on the project (e.g., event attributes, JSON fields)

The exact data types depend on the team and engagement.

14. Will I participate in metric/KPI definition and validation?

Yes. A key part of the role is supporting:

- Metric definition (precise logic, scope, filters, time windows, inclusion/exclusion rules)
- Calculation and validation (reconciliation, edge cases, consistency over time)
- Documentation so metrics are consistent across teams and reusable in reporting

This is often where interns add significant value quickly.

15. Does the role involve interaction with product/business/tech teams?

Yes. The internship usually involves collaboration to:

- Understand the problem context and decision needs
- Align hypotheses and definitions
- Interpret results and constraints
- Ensure analyses are actionable and correctly understood

You'll typically interact through your project team and learn how to translate between data and business questions.

16. How does the work model operate (remote/hybrid), expected working hours, and flexibility?

The work model is commonly **remote**, with frequent collaboration through digital tools.

Working hours generally align with team needs and project timelines. When communicated early, academic constraints (classes/exams) are typically manageable through planning and clear expectations around deliverables.

17. How does the selection process work?

A typical process includes:

- A logic/quant-style test (maybe GMAT-like depending on the role)
- A basic technical case (Python/SQL fundamentals + analytical reasoning)
- A conversation focused on problem-solving approach, communication, and fit

Exact steps may vary by hiring cycle.

18. Is there a possibility of full-time employment after the internship?

Yes. Internships are often an entry point to junior roles, depending on:

- Performance during the internship
- Business demand and project pipeline - available openings

Example Project

Product Profitability Deep-Dive (B2B Beverage Importer — Brazil)

Context

A Brazilian B2B company that imports beverages (multiple brands, categories, and packaging formats) wanted to understand true product profitability beyond gross margin. The leadership team faced recurring questions like:

- “Which SKUs are actually profitable after freight, taxes, and trade terms?”
- “Are we making money on certain customers or channels, or are discounts and logistics wiping it out?”
- “Which cost drivers explain margin erosion month-over-month?”

The company had data spread across ERP exports, freight invoices, tax components, pricing tables, and sales/order systems, and needed a single, consistent profitability view with drill-down capability.

Objective

Build a structured profitability model that allows analysis at multiple breakdowns (e.g., SKU → brand → category, customer → segment, channel, region, warehouse) and deliver:

- a standardized profitability dataset (single source of truth)
- a profitability waterfall showing margin bridges and cost-to-serve
- a Power BI dashboard enabling interactive drill-down and decision-ready views

What the intern would do (with supervision)

1. Data mapping and dataset assembly (Python)

The work starts by mapping the main data sources across the company’s commercial and financial flow—sales orders/invoices, import and landed cost components (FOB/CIF and fees), freight and last-mile distribution, tax/fiscal fields, discounts and commercial agreements, and operational adjustments like returns and credit notes. From there, we build a clean, consistent dataset at the ****invoice line / SKU level****, with standardized keys (SKU/brand/category, customer/segment/channel/region, time period, and warehouse/DC). Throughout the build, we run basic sanity checks and reconciliations to ensure totals match finance references (revenue, COGS, discounts) and to flag gaps such as missing cost allocations or unusual unit price/cost outliers.

2. Profitability model design (definitions + allocation rules)

Next, we implement a profitability logic that goes beyond gross margin by structuring the P&L into clear layers: revenue (gross revenue → discounts/promos → rebates/trade terms → net revenue), cost of goods including import structure (base purchase cost plus duties/fees, port/customs/handling, and any tracked FX adjustments), and cost-to-serve (freight to warehouse, distribution to customer, warehousing/handling, commissions/channel fees, and returns/breakage). When costs are not naturally available at SKU level, we apply transparent allocation rules (e.g., freight by weight/volume, warehousing by pallet-days, overhead by revenue share). The priority is making assumptions explicit, consistent, and traceable so the model can be explained and trusted.

3. Build the profitability waterfall (margin bridge)

With the profitability layers defined, we build a standardized waterfall that explains how you move from top-line revenue down to contribution margin, and we make it drillable across levels such as company, brand, category, SKU, customer segment, and channel. The waterfall typically bridges from Gross Revenue to Net Revenue, then subtracts COGS and import components to reach Landed Gross Margin, and finally subtracts logistics, warehousing, commercial costs, and returns to arrive at a cost-to-serve adjusted Contribution Margin. A key part of the work is ensuring the waterfall ties out mathematically at each drill level and is easy to narrate to stakeholders.

4. Power BI dashboard (interactive profitability cockpit)

Finally, we translate the model into a Power BI dashboard that allows stakeholders to explore profitability quickly and consistently—by SKU/brand/category, customer/segment/channel, region/warehouse/route, and winner/loser analysis with clear drivers (discount intensity, freight, taxes, returns, etc.). The dashboard typically includes an executive summary view, a dynamic waterfall page, deep dives for SKU and customer/channel profitability, and a driver analysis section comparing changes over time. On the backend, this usually requires a clean star-schema-like model, well-defined measures for each margin layer and allocation, and validation views that reconcile dashboard numbers to finance totals.

Typical deliverables

- A documented, versioned profitability dataset (invoice-line level) with clear definitions
- A standardized waterfall logic (consistent margin layers and driver components)
- A Power BI dashboard with drill-down and export-ready tables
- A short written summary highlighting key insights (e.g., “Top 10 loss-making SKUs and why”)

Example insights this project enables

- Identifying SKUs with strong gross margin but negative contribution margin due to distribution cost
- Detecting customers where discounts + payment terms erase profitability
- Understanding whether margin erosion is driven by price/mix vs landed cost vs logistics
- Prioritizing renegotiation targets (freight contracts, trade terms, minimum order quantities)

Skills used

Python, SQL, data cleaning, metric definition, allocation logic, reconciliation/QA, visualization design, Power BI (data model + measures), and structured communication of insights.