Universitat Pompeu Fabra Barcelona

UPF – Gestió de projectes

Deliverable: 1.4.1 — Decision & Assumptions Log

AIGÜES DE BARCELONA - INCIDÈNCIES EN COMPTADORS INTEL·LIGENTS

Deliverable: 1.4.1 — Decision & Assumptions Log
Project: *Incidències en Comptadors Intel·ligents (Aigües de Barcelona)*
Team: 102F
Project Manager (PM): Arnau Rodon Comas
Project Core Team (PCT): Arnau Rodon, Asul Garcia, Jordi Esteve, Albert Jané, Nahia Anaut

Iteration: W1–W10 (all)
Version: v1.0
Date: 30/11/2025

This log consolidates the main **decisions** and **assumptions** taken during the project, with a focus on:

- **Data readiness & integration** (Iteration 1; WBS 1.1)
- **Statistical analysis and model baselines** (Iteration 2; WBS 1.2.1–1.2.2)
- **Evaluation, labeling & forecasting design** (WBS 1.2.3)
- **Project management & governance** (WBS 1.4.x)

It supports **Rq5 – Reproducibility** and traceability requirements defined in the Project Work Plan.

1. Log Structure

Each entry is tagged as:

- **Type**: Decision (D) or Assumption (A)
- **Area**: Data / Modeling / Evaluation / PM–Governance / Environment
- **Impact**: Scope / Quality / Schedule / Risk
- **Status**: Confirmed (C), Provisional (P), Open (O)
- **Links**: Main artefacts where the decision/assumption is implemented or referenced.

2. Log Structure

2.1 Data readiness, cleaning & integration

ID	Type	Area	Description	Rationale	Impact	Status	Links
D-01	Decision	Data	Filter to records with FECHA ≥ DATA_INST_COMP. Rows where the consumption date precedes the installation date are dropped.	Rows with missing operational metadata and inconsistent dates are interpreted as pre-installation artefacts, not valid measurements. 1.1.1.b_Validation_Report (1)	Quality: Removes structurally invalid readings; Scope: Reduces volume but increases trust in remaining data.	C	1.1.1.b Validation Report; Unified dataset v1 (1.1.1) 1.1.1.b_Validation_Report (1)
D-02	Decision	Data	Deduplicate on (POLIZA_SUMINISTRO, FECHA), keeping a single record when all attributes match.	Duplicate rows with identical attributes are treated as pure replications and are removed to avoid double-counting. 1.1.1.b_Validation_Report (1)	Quality: Eliminates bias in counts and statistics; Scope: No information loss, only redundancy removal.	C	1.1.1.b Validation Report 1.1.1.b_Validation_Report (1)

D-03	Decision	Data	Cast DIAM_COMP, CODI_MODEL, NUM_MUN_SGAB, NUM_DTE_MUNI to categorical types (not numeric).	Values represent discrete classes (diameter codes, model ids, municipality codes) rather than continuous quantities; treating them as categories avoids misleading numeric operations. 1.1.1.b_Validation_Report (1)	Quality: Correct semantics for modeling; Scope: Enables safer encodings and grouping.	C	1.1.1.b Validation Report; Feature-engineering notebook (1.2.2.a) 1.1.1.b_Validation_Report (1)
D-04	Decision	Data	Standardise FECHA to datetime and align with weather data on (FECHA, NUM_MUN_SGAB).	Ensures robust temporal joins and allows consistent integration of weather at municipality level. 1.1.1.b_Validation_Report (1)	Quality: Correct calendar alignment; Scope: Enables joint analysis of consumption and weather.	C	1.1.1.b Validation Report; Data integration scripts 1.1.1.b_Validation_Report (1)
D-05	Decision	Data	Impute missing precipitation as 0 ("no rain") after merge.	Weather provider uses missing precipitation as "no recorded rain"; imputing zeros maintains analytical coherence. 1.1.1.b_Validation_Report (1)	Quality: Avoids artificial gaps; Risk: Misclassification if source semantics change (mitigated by documentation).	C	1.1.1.b Validation Report 1.1.1.b_Validation_Report (1)
A-01	Assumption	Data	Dataset is representative of business behaviour and anomaly patterns despite being academic and anonymised.	Work Plan states that datasets are assumed representative for analysis and modeling; no additional sampling design is available. Team102F.ProjectWorkPlan.v3.1	Risk: If not representative, model performance and anomaly rates may not generalise to production.	P	Project Work Plan – Project Constraints / Assumptions Team102F.ProjectWorkPlan.v3.1

A-02	Assumption	Data	High missingness in operational fields for pre-installation rows implies meters not yet in service.	Cleaning strategy and temporal filter are built on this interpretation; no contradictory business rule was provided. 1.1.1.b_Validation_Report (1)	Quality: Justifies dropping those rows; residual risk is documented in data-quality notes.	C	1.1.1.b Validation Report; Data Quality Log (1.4.3) 1.1.1.b_Validation_Report (1)

2.2 Statistical analysis & feature engineering

ID	Type	Area	Description	Rationale	Impact	Status	Links
D-06	Decision	Modeling – Features	Use log(1 + CONSUMO_REAL) as the canonical level feature (CONSUMO_LOG).	Raw consumption is zero-inflated and heavily right-skewed; log transform yields a more symmetric distribution and stabilises variance.	Quality: Improves robustness of clustering and LSTM training; Scope: All downstream models operate on log-transformed levels.	C	1.2.1 Statistical Analysis; 1.2.2 Model baselines
D-07	Decision	Modeling – Features	Compute both classical and robust z-scores per meter (mean/σ and median/MAD).	Different meters have heterogeneous distributions; robust metrics better capture anomaly intensity in presence of outliers. 1.2.2_Model_Baselines	Quality: Enables richer anomaly-intensity features; Scope: Robust z-scores primarily used in LSTM feature space, not in k-means.	C	Feature set & engineering notebook (1.2.2.a) 1.2.2_Model_Baselines
D-08	Decision	Modeling – Features	Encode temporal structure via calendar features (day of week, month, season, weekend) plus sine/cosine cyclical encodings.	Weekly and seasonal patterns are visible; cyclical encodings avoid artificial distance between December and January, etc. 1.2.2_Model_Baselines	Quality: Improves separation of consumption regimes (weekday/weekend, summer/winter); supports both k-means and LSTM.	C	1.2.2 Model baselines; Feature engineering code 1.2.2_Model_Baselines

D-09	Decision	Modeling – Features	Use rolling statistics over 7, 14, 30 and 90 days per meter (means, std, medians, CV, deviations) with minimum history constraint.	Rolling windows capture local baselines and volatility at multiple horizons; history minimum avoids unstable statistics at meter start-up. 1.2.2_Model_Baselines	Quality: Key for anomaly detection; Scope: Reduced subset sent to k-means; full set used for LSTM.	C	1.2.1 Statistical Analysis; 1.2.2 Model baselines
D-10	Decision	Modeling – Features	Reconstruct numeric DIAMETER and derive DIAMETER_RISK, retain model/brand flags and NUM_MODELS.	Statistical analysis shows diameter and meter type strongly influence consumption and anomaly rates; derived risk score reflects observed patterns.	Quality: Encodes hardware-driven risk; Scope: Used across both k-means and LSTM to capture product effects.	C	1.2.1 Statistical Analysis; 1.2.2 Model baselines
D-11	Decision	Modeling – Features	Create anomaly-proxy features (IS_ZERO_CONSUMPTION, rolling percentiles, MACD, SUDDEN_CHANGE_SCORE) but exclude them from k-means features.	These derived features are useful as signals for supervised learning but would create circularity if they drove unsupervised clusters whose output defines anomalies. 1.2.2_Model_Baselines	Quality: Keeps k-means focused on baseline behaviour; Scope: Proxies are only used as LSTM inputs and descriptive diagnostics.	C	Feature engineering notebook (1.2.2.a); Model baselines 1.2.2_Model_Baselines
A-03	Assumption	Modeling – Features	Weather is a weak short-term driver of consumption and anomalies; its role is mainly contextual.	Correlation analysis shows negligible linear relationship between daily weather and consumption; seasonal usage is already captured via calendar features. 1.2.1_Statistical_Analysis (1)	Scope: Weather kept in LSTM feature space but is not a primary driver; Risk: Underestimates complex weather–usage interactions (consider for future work).	C	Statistical Analysis Report (1.2.1); Model baselines

2.3 Anomaly definition & k-means layer

ID	Type	Area	Description	Rationale	Impact	Status	Links
A-04	Assumption	Evaluation	No authoritative anomaly labels exist; anomalies must be inferred algorithmically.	Status Report highlights the lack of an approved anomaly labeling policy; current work uses proxy labels and thresholds. 102F.ProjectStatusReport.v2.0	Risk: Metrics (Recall/FPR) are conditional on proxy labels; confidence in absolute values is limited until business rules are defined.	O	Project Status Report v2.0; Evaluation pack (1.2.3) 102F.ProjectStatusReport.v2.0
D-12	Decision	Modeling – Anomaly detection	Adopt a two-layer architecture: per-brand k-means for static anomaly labeling + per-brand LSTM for anomaly risk forecasting.	Enables unsupervised definition of anomalies from structure in the data, then supervised learning of future risk based on those labels. 1.2.2_Model_Baselines	Scope: Clear separation between labeling and forecasting; Quality: Uses complementary strengths of unsupervised and supervised methods.	C	1.2.2 Model baselines; Intro to Water-Anomaly Detection 1.2.2_Model_Baselines
D-13	Decision	Modeling – Anomaly detection	Partition data by meter BRAND and train a separate k-means model per brand (with minimum record thresholds).	Brand-specific behaviour and deployment volumes differ; per-brand models avoid large brands dominating cluster structure. 1.2.2_Model_Baselines	Quality: Better captures hardware-specific patterns; Scope: Requires per-brand model management.	C	Model baselines – K-means module 1.2.2_Model_Baselines

D-14	Decision	Modeling – Anomaly detection	Use RobustScaler on the k-means feature subset and MiniBatchKMeans for scalable clustering.	RobustScaler mitigates the impact of outliers; MiniBatchKMeans scales to millions of records while remaining reproducible. 1.2.2_Model_Baselines	Quality: Stable centroids despite heavy tails; Schedule: Reduces runtime; Risk: Approximate optimisation vs full KMeans deemed acceptable.	C	Model baselines – K-means design 1.2.2_Model_Baselines
D-15	Decision	Modeling – Anomaly detection	Optimise k (clusters) in a brand-specific [min_clusters, max_clusters] range using a combined metric (Silhouette, DB, CH, inertia) with tie-breaking in favour of smaller k.	Balances cluster separation and compactness while preferring simpler models when quality differences are marginal. 1.2.2_Model_Baselines	Quality: Avoids overfitted cluster structures; Scope: Adds a controlled optimisation loop per brand.	C	Model baselines – K-means optimisation 1.2.2_Model_Baselines
D-16	Decision	Modeling – Anomaly detection	Support both Euclidean and Mahalanobis distances, with regularised covariance and fallback to Euclidean if ill-conditioned.	Mahalanobis accounts for covariance structure within clusters; regularisation and fallbacks ensure numerical stability. 1.2.2_Model_Baselines	Quality: More faithful distance in well-behaved clusters; Risk: Complexity mitigated by robust fallbacks.	C	K-means implementation; Model baselines 1.2.2_Model_Baselines
D-17	Decision	Evaluation – Static anomalies	Define anomalies as the top α tail (currently ≈2.5%) of cluster distances per cluster; compute anomaly_score as distance/threshold.	Tail-quantile approach yields controlled and comparable anomaly rates per cluster, independent of cluster-specific variance. 1.2.2_Model_Baselines	Quality: Consistent anomaly rate; Scope: Global anomaly rate dictated by α; Risk: Requires domain validation of α value.	P	K-means configuration; Evaluation pack – anomaly rate analysis 1.2.2_Model_Baselines

D-18	Decision	Environment / Reproducibility	Use fixed random_state and capped subsampling for k-means training; persist per-brand models, scalers and thresholds.	Ensures that repeated runs with the same configuration produce identical labels and can be reproduced at delivery time. 1.2.2_Model_Baselines	Quality & Rq5: Strong traceability and reproducibility; Scope: Additional metadata and storage for configuration and models.	C	Model baselines; Environment spec & Runbook (1.4.2)

2.4 LSTM forecasting layer & evaluation

ID	Type	Area	Description	Rationale	Impact	Status	Links
D-19	Decision	Modeling – Forecasting	Use sliding 90-day sequences per meter as LSTM inputs, predicting anomaly occurrence at 3 horizons: next day, next 7 days, next 30 days.	90 days is long enough to capture seasonal and trend patterns while keeping sequences tractable; multiple horizons reflect operational needs. 1.2.2_Model_Baselines	Scope: Three parallel binary targets per sequence; Quality: Captures both short- and medium-term risk.	C	Model baselines – LSTM section; 1.2.2 Training runs 1.2.2_Model_Baselines
D-20	Decision	Modeling – Forecasting	Train one LSTM model per brand, with train/validation splits at meter level (no meter appears in both).	Prevents leakage from overlapping windows of the same meter, improving generalisation estimates; mirrors per-brand design of k-means. 1.2.2_Model_Baselines	Quality: More realistic validation; Scope: More models to train and maintain.	C	Model baselines – LSTM; 1.2.2.b Training & tuning 1.2.2_Model_Baselines

D-21	Decision	Modeling – Forecasting	Input feature space for LSTM = full engineered features + k-means outputs (cluster, distance, is_anomaly, anomaly_score).	Allows the LSTM to learn how sequences of anomaly intensity, cluster regimes and contextual variables evolve before future anomalies. 1.2.2_Model_Baselines	Quality: Richer temporal representation; Scope: High-dimensional inputs increase model complexity but are justified by strong regularisation.	C	Feature engineering (1.2.2.a); LSTM datasets 1.2.2_Model_Baselines
D-22	Decision	Modeling – Forecasting	Architecture: two stacked LSTM layers + dense head with dropout and batch norm; 3 sigmoid outputs (one per horizon).	Standard yet expressive architecture for multihorizon binary risk prediction; aligns with limited data and compute. 1.2.2_Model_Baselines	Quality: Balance between capacity and overfitting control; Scope: Enables per-horizon probabilities for later thresholding and risk mapping.	C	Model baselines – LSTM design 1.2.2_Model_Baselines
D-23	Decision	Evaluation	Address class imbalance via per-sample, per-horizon weights; give higher weight to positives and to shorter horizons (day > week > month).	Positive anomaly sequences are rare; without weighting the network would under-predict anomalies. Short-horizon alerts are operationally more critical. 1.2.2_Model_Baselines	Quality: Increases Recall on rare events while controlling FPR; Risk: Requires careful threshold setting.	C	Evaluation pack (1.2.3.a Metrics report)

A-05	Assumption	Evaluation	Target bounds on FPR and Recall (Rq3: ≤ X% FPR, ≥ Y% Recall) are business-reasonable but not yet approved by a real PO.	Work Plan defines these targets generically; actual X, Y values and business tolerance are not formally validated with Aigües de Barcelona. Team102F.ProjectWorkPlan.v3.1	Risk: Achieving nominal targets may not align with real-world utility; documented explicitly in metrics report.	O	Requirements Rq3; 1.2.3.a Metrics report; Status Report v2.0
D-24	Decision	Evaluation / Reporting	Expose qualitative risk levels (LOW / MEDIUM / HIGH / CRITICAL) derived from the maximum predicted probability across horizons and configured thresholds.	Facilitates interpretation by non-technical stakeholders and supports prioritisation of field inspections. 1.2.2_Model_Baselines	Scope: Adds a post-processing layer; Quality: Risk mapping is transparent and configurable in the runbook.	C	LSTM prediction pipeline; Visualization pack (1.3.2) 1.2.2_Model_Baselines

2.5 Project management, governance & environment

ID	Type	Area	Description	Rationale	Impact	Status	Links
D-25	Decision	PM – Schedule	Re-baseline Iteration 2 (modeling) around actual dataset delivery on 28/10, pulling internal deadlines 2–3 days earlier per milestone.	Original plan assumed earlier data availability; re-baselining protects official due dates and avoids last-minute compression.	Schedule: Restores feasibility; Risk: Documented as Issue 01 (now closed).	C	Project Work Plan v3.1; Project Status Report v2.0 – change 01

D-26	Decision	PM – Governance	Treat the Project Owner as a conceptual role; PSC (Solution Providers) emulate PO decisions; PM is not part of PSC and attends by invitation.	The challenge is academic; no active client-side PO. Clarifying roles avoids ambiguity in who approves scope and artefacts. Team102F.ProjectWorkPlan.v3.1	Quality & Traceability: Clear accountability for approvals and escalations; logged in Work Plan v3.1.	C	Project Work Plan – Stakeholders & RACI Team102F.ProjectWorkPlan.v3.1
D-27	Decision	PM – Comms	Adopt governance cadence: weekly PM↔SP sync (15’), monthly PSC review (45’), with decisions logged in this log (1.4.1).	Ensures regular alignment, timely escalation, and traceable decisions as required by methodology. Team102F.ProjectWorkPlan.v3.1	Schedule: Supports early detection of issues; Quality: Better-controlled scope and risk.	C	Project Work Plan – Communications Management; Status Reports
D-28	Decision	Environment / CM	Use Google Drive as the controlled document repository and GitHub for code & notebooks, with standard naming (Group102F.DocName.vX.X) and commit hashes referenced in documents.	Ensures configuration management, traceability from documents to code, and single source of truth for controlled artefacts. Team102F.ProjectWorkPlan.v3.1	Quality & Rq5: Enables full reproducibility and auditability; Scope: Requires discipline in versioning and documentation.	C	Configuration Management section; Environment spec & Runbook (1.4.2) Team102F.ProjectWorkPlan.v3.1
A-06	Assumption	PM – Scope	Out of scope: production deployment, live dashboards for external users, and real-time alerting; project delivers an analytical prototype and documentation only.	Stated explicitly in Work Plan as out-of-scope items given academic constraints and limited compute. Team102F.ProjectWorkPlan.v3.1	Scope: Prevents scope creep into MLOps and real-time systems; Risk: Prototype must be clearly framed as non-production.	C	Project Work Plan – Constraints / Out of Scope Team102F.ProjectWorkPlan.v3.1

3. Maintenance of Log

Per WBS component description for WP 1.4.1 – Decision & Assumptions Log, this document should be:

- Updated weekly by the PM based on meeting minutes and PSC decisions.
  Team102F.ProjectWorkPlan.v3.1

- Cross-referenced from key deliverables (Validation Report 1.1.1.b, Statistical Analysis 1.2.1, Model Baselines 1.2.2, Evaluation Pack 1.2.3, Final Report 1.3.1).

- Stored under configuration management in Google Drive with versioning (e.g., 1.4.1_DecisionAssumptionLog.docx).