

SICRETO LABS LLC

Sample Task Pack: 10 Expert Tasks for Spain- Context Hospitality Agents

Paper sample showing task format, difficulty and grading approach

Version: paper sample v0.1 · 2026-10-10

Prepared by: Ramon Sicre, Sicreto Labs LLC, Andalucía, Spain

Contents: method note · format of each task · index · tasks H01–H10 · provenance notes

All hotels, restaurants, companies, people, tax IDs and reservation numbers in this pack are fictional.
This is a sample for evaluating format and difficulty. It is not legal, tax or labour advice.

Method note

- **What this is:** 10 example tasks showing the format, difficulty and grading approach we would deliver as expert tasks and inside simulated hospitality environments (A1 hotel, A2 restaurant & events, A3 club & experiences concierge). It is a paper sample: no environment is running yet, no model has been tested on these tasks, and no paid expert has reviewed them.
- **How it was made:** written from general professional knowledge of Spanish hospitality operations. No capture, no client files, no real guest or company data.
- **Everything named is fictional** and marked (*ficticio*) where first used. OTA names are placeholders ("OTA-X", "OTA-Y").
- **Rules and rates.** Where a task depends on a legal, tax or labour rule, the rule or rate is stated inside the task as a scenario assumption, so the task tests applying it rather than recalling it. These assumptions are not legal advice and have not been verified by a licensed professional. Before any task is used as a gold reference, a Spain-qualified professional (gestor/asesor fiscal, labour adviser, tourism or food-safety specialist, as relevant) should confirm them.
- **Language.** Prompts are in Spanish (Spain), with an English summary. Production tasks can ship ES only, EN only or both.
- **Likely failure modes** are hypotheses for the task designer. They are not results from testing any model.
- **Grading.** Rubrics sum to 100. **[P]** = checkable programmatically (state diff / numeric / exact match). **[R]** = judged by a human or LLM grader against the gold key.

Format of each task

Field	Content
ID / role / difficulty	Task ID, the professional role the agent plays, difficulty tier (T1 easy – T3 hard)
Prompt	What the agent receives (ES) + EN summary
Context	Self-contained data the agent needs (fictional)
Expected outputs	The artefacts the agent must produce
Gold answer (key)	The reference answer or the required elements
Rubric	Weighted items (sum = 100). [P] = checkable programmatically (state diff / numeric / exact match). [R] = rubric judged by a human or LLM grader against the gold key
QA notes	Review notes, tolerances, likely failure modes (hypotheses)

Index

#	Task	Role	Tier	Maps to
H01	Overbooking: who to walk and how	Front-office manager	T2	A1 Hotel PMS + channel manager
H02	Group block displacement analysis	Revenue manager	T3	A1
H03	Traveller registration audit	Front-office / compliance	T2	A1 (simulated traveller portal (simulated))
H04	Corporate factura completa, incl. replacing a factura simplificada	Night auditor / billing	T2	A1
H05	Allergen matrix and safe menu for a coeliac + sesame-allergic guest	F&B manager / chef	T2	A2 Restaurant & events ops
H06	Wedding quote, deposit schedule and minimum-guarantee recalculation	Events manager	T3	A2
H07	Staff rota compliance check and fix	Restaurant manager / HR	T3	A2
H08	Andalusian wine pairing for a 5-course tasting menu	Sommelier	T2	A2 / A3
H09	Golf + padel + winery itinerary with weather rescheduling	Concierge / DMC planner	T3	A3 Club & experiences concierge
H10	Non-refundable OTA booking dispute + formal complaint request	Guest relations manager	T2	A1 / A3

H01 — Overbooking: who to walk and how

Role: Front-office manager, 4-star hotel · **Tier:** T2 · **Environment:** A1

Prompt (ES):

Son las 12:00. Por un fallo de sincronización del channel manager, el Hotel Mirador del Puerto (*ficticio*), 4 estrellas, 80 habitaciones, en la provincia de Málaga, tiene 82 reservas confirmadas para esta noche y 0 habitaciones fuera de servicio recuperables. Decide qué 2 reservas reubicar, justifica la elección, redacta el mensaje al huésped (ES o EN según el idioma de la reserva), indica qué costes asume el hotel y qué acciones correctivas harías hoy en el channel manager y en el PMS.

EN summary: The hotel is 2 rooms oversold tonight because of a channel-manager sync failure. Choose which 2 reservations to relocate ("walk"), justify the guest message, state the costs the hotel covers, and list the corrective actions in the channel manager and PMS.

Context (fictional):

Res.	Channel	Nights	Rate	ETA	Guest profile / notes
R01	Direct	3	Flexible	16:00	Loyalty Gold, anniversary noted (cava ordered)
R02	OTA-X	1	Non-refundable	23:30	Solo business traveller, first stay, EN
R03	OTA-Y	1	Flexible	15:00	Family with 2 children, connecting rooms requested
R04	Corporate (Distribuciones Ejemplo S.L., ficticio)	2	Corporate	18:00	Account with 120 room-nights/year
R05	OTA-X	1	Flexible	22:00	Couple, first stay, no requests, ES
R06	Direct (group block)	5	Group	14:00	Wedding guest, part of a 12-room block
R07	OTA-Y	1	Flexible	19:00	Guest with reduced mobility, accessible room requested
R08	Direct	1	Flexible	21:00	Repeat guest, 6 previous stays
R09–R82	Mixed	2–7	Mixed	Mixed	No special flags

- Partner hotel: Hotel Brisa (*ficticio*), 4 stars, 600 m away, 3 standard rooms available tonight at €130, **no accessible rooms available**.
- Hotel policy (*ficticio*): relocate to equal or higher category nearby, hotel pays the night and transport both ways, offer to bring multi-night guests back the next day, priority protection for loyalty, groups, corporate accounts, accessibility needs and families with children.

Expected outputs: (1) the 2 reservation IDs; (2) a short justification against walk criteria; (3) guest message; (4) cost list; (5) corrective actions.

Gold answer (key):

1. Walk **R02 and R05**: 1-night stays, first-time guests, late arrival, no special needs, OTA bookings with no loyalty or account value. Both fit the partner hotel (standard rooms, same category).
2. Must not walk: R01 (loyalty + occasion), R03 (family, connecting rooms), R04 (corporate account), R06 (group block, multi-night), R07 (accessibility need; partner has no accessible room), R08 (repeat guest).
3. Contact the guests **before arrival** (phone + email/OTA message), apologise, explain plainly, confirm the partner hotel address and that the hotel pays the room and taxi. Add a goodwill gesture (e.g. voucher or free night on a future stay) per policy. Under the fictional hotel policy, R02's non-refundable rate does not change the hotel's commitment to provide accommodation.

4. Costs: 2 × €130 at the partner hotel + taxis; the guest pays nothing beyond what they had already agreed to pay. The original booking is settled per the hotel's policy (e.g. no charge to the guest, or charge and pay the partner, but the guest is never worse off).
5. Corrective actions: close availability on all channels for tonight (stop-sell), reconcile inventory between PMS and channel manager, find the sync error (mapping / rate plan / allotment), check the next 14 days for the same error, notify the OTA (relocation note on R02/R05), log the incident, and record the relocation in the PMS so housekeeping and night audit figures are correct.

Rubric (100):

#	Item	Pts	Type
1	Selects exactly R02 and R05	25	[P]
2	Does not select R01/R03/R04/R06/R07/R08, and gives the reason for at least 4 of them	10	[R]
3	Relocation to equal/higher category nearby, hotel pays room and transport	15	[R]
4	Contacts guests before arrival; message is clear, apologetic, correct language, includes address and what is covered	15	[R]
5	States the non-refundable rate does not remove the hotel's commitment under its policy	5	[R]
6	Stop-sell on all channels + PMS/channel manager reconciliation	15	[P]/[R] (in A1: state diff shows availability closed and inventory matching)
7	Checks forward dates for the same error	10	[R]
8	Logs incident / notifies OTA	5	[R]

QA notes: The task relies on the fictional hotel policy given in the context, not on any statutory relocation rule. Likely failure modes (hypothesis): walking the accessible-room or family booking because it is "only 1 night"; forgetting the stop-sell; treating the non-refundable guest as having fewer rights.

H02 — Group block displacement analysis

Role: Revenue manager · Tier: T3 · Environment: A1

Prompt (ES):

Un organizador de bodas pide un bloqueo de 40 habitaciones durante 3 noches (N1, N2, N3) a 160 € netos por habitación y noche, más un banquete de 80 personas a 70 € por persona la noche N2. ¿Aceptas, rechazas o haces una contraoferta? Calcula el beneficio neto incremental (contribución) frente a la demanda transitoria prevista, por noche y total, y si contraofertas, calcula la tarifa mínima de N2 para no perder dinero esa noche. Todos los importes sin IVA.

EN summary: Accept, reject or counter a 40-room × 3-night wedding block at €160 net plus an 80-person banquet at €70/person on night 2. Compute incremental contribution vs forecast transient demand, per night and total, and the break-even night-2 group rate.

Context (fictional, Hotel Mirador del Puerto):

- Capacity: 120 rooms.
- Unconstrained transient demand forecast: N1 = 80 rooms, N2 = 118, N3 = 90.
- Transient ADR (net): N1 €150, N2 €220, N3 €170.
- Cost per occupied room (CPOR): €25.
- Transient ancillary spend per occupied room: €30 at 40% margin (= €12 contribution).
- Banquet F&B margin: 35%. Group ancillary beyond the banquet: ignore.
- With the group, rooms available to transient = 120 – 40 = 80.

Expected outputs: displaced rooms per night; contribution table; recommendation; break-even N2 rate; suggested contract protections.

Gold answer (key):

Night	Displaced transient rooms	Group rooms contribution $40 \times (160 - 25)$	Banquet contribution	Displaced contribution (rooms + ancillary)	Net
N1	0 (80 – 80)	€5,400	—	€0	+ €5,400
N2	38 (118 – 80)	€5,400	$80 \times 70 \times 35\% =$ €1,960	$38 \times 195 = €7,410 + 38 \times 12 =$ €456 → €7,866	–€506
N3	10 (90 – 80)	€5,400	—	$10 \times 145 = €1,450 + 10 \times 12 =$ €120 → €1,570	+ €3,830
Total	48	€16,200	€1,960	€10,436	+ €8,724

- Recommendation: accept overall (positive €8,724) but **counter on N2**: N2 is loss-making at €160.
- Break-even N2 group rate: $40 \times (r - 25) + 1,960 \geq 7,866 \rightarrow r \geq \mathbf{€172.65}$ (round to $\geq \mathbf{€173}$). Alternatives: reduce the N2 block (e.g. to fewer rooms), or keep €160 and require a higher banquet minimum spend.
- Protections: attrition clause (e.g. allowed reduction % and dates), cut-off date for releasing unpicked rooms, rooming list deadline, deposit schedule, and a forecast-risk note (transient demand is a forecast; wash factor).

Rubric (100):

#	Item	Pts	Type
1	Displaced rooms 0 / 38 / 10	15	[P]
2	Night nets +5,400 / –506 / +3,830 (±€1)	25	[P]
3	Total +€8,724 (±€1)	10	[P]
4	Includes CPOR and displaced ancillary (not revenue only)	10	[R]

#	Item	Pts	Type
5	Break-even N2 rate €172.65 (accept €172.6–€173)	15	[P]
6	Recommendation = accept with N2 counter (or equivalent N2 fix)	15	[R]
7	At least 3 sensible contract protections	10	[R]

QA notes: Pure arithmetic on given data, no external facts. Likely failure modes (hypothesis): comparing revenue instead of contribution; counting displacement on N1; using capacity 120 instead of 80 for transient; forgetting displaced ancillary.

H03 — Traveller registration audit

Role: Front-office supervisor / compliance · **Tier:** T2 · **Environment:** A1 with a *simulated* traveller-communication portal (no connection to any real government system)

Prompt (ES):

Revisa los 6 registros de entrada de ayer del Hotel Mirador del Puerto (*ficticio*) antes de enviarlos al portal de comunicación de viajeros (simulado). Para cada registro, indica qué dato obligatorio falta o es incorrecto y si la comunicación está fuera de plazo. Después, propone el procedimiento para corregirlo y evitar que se repita.

EN summary: Audit 6 check-in records against the scenario rules below before submitting to a simulated portal; list missing/wrong data and late communications; propose a fix procedure.

Scenario rules (assumed for this task, not legal advice): each record must carry the document type, number and support number where the document has one, nationality, date of birth, full address, phone and/or email, payment data, and, for a minor, the relationship to the accompanying adult. Records must be communicated within 24 hours of check-in.

Context (fictional records; "—" = empty field):

Rec	Traveller	Doc type / number	Doc support no.	Nationality	DOB	Address	Phone / email	Payment data	Check-in → communicated
C1	Adult, ES	DNI 00000000T	—	ES	1980-04-02	Complete	Both	Card, recorded	09/10 15:00 → 09/10 18:00
C2	Adult, DE	Passport X0000000	n/a	DE	1975-11-20	Complete	Both	Card, recorded	09/10 16:00 → 09/10 20:00
C3	Minor (14), travelling with C2	Passport X0000001	n/a	DE	2012-06-01	Complete	— (minor)	n/a	09/10 16:00 → 09/10 20:00 · relationship to adult: —
C4	Adult, FR	ID card 000000000000	—	FR	—	Complete	Both	Card, recorded	09/10 17:30 → 09/10 19:00
C5	Adult, GB	Passport 000000000	n/a	GB	1990-01-15	City only	Phone only	Card, recorded	09/10 21:00 → 09/10 22:00
C6	Adult, ES	DNI 00000001R	Recorded	ES	1988-08-08	Complete	Both	—	08/10 13:00 → 10/10 09:00

Expected outputs: a per-record issue list; late-communication flag(s); correction procedure; prevention measures.

Gold answer (key):

- C1: missing **document support number** (número de soporte) for the Spanish DNI.
- C2: no issues.
- C3: minor's **relationship (parentesco)** with the accompanying adult missing.
- C4: missing **date of birth**; check whether the support number applies to that ID document.
- C5: **incomplete address** (full usual residence address required); email missing (record what is available; check current rules on which contact fields are mandatory).
- C6: **payment data missing**, and communication **outside the deadline** (more than 24 h after check-in, per the scenario rule).
- Procedure: complete the data with the guest (at desk / by message), submit corrections, document the late submission and reason, keep records for the retention period that applies, and do not collect more than the scenario rule requires (data minimisation).
- Prevention: make the mandatory fields required in the PMS check-in form, add a 24-hour submission alert, and run a daily audit at night audit.

Rubric (100):

#	Item	Pts	Type
1	C1 support number	10	[P]
2	C2 correctly marked OK	5	[P]
3	C3 relationship for the minor	15	[P]
4	C4 date of birth	10	[P]
5	C5 incomplete address	10	[P]
6	C6 payment data missing	10	[P]
7	C6 late communication (>24 h)	15	[P]
8	Correction procedure incl. documenting late submission	10	[R]
9	Prevention (required fields, alert, daily audit)	10	[R]
10	Mentions retention / data minimisation sensibly	5	[R]

QA notes: The scenario rules above are task assumptions. A Spain-qualified compliance professional should confirm the real field list, deadline and retention period before this task is used as a gold reference. The portal is a simulation and must never connect to or imitate a real government site or its branding. Likely failure modes (hypothesis): missing the minor's relationship field; not spotting the late timestamp; asking for data the scenario does not require.

H04 — Corporate factura completa, incl. replacing a factura simplificada

Role: Night auditor / billing · **Tier:** T2 · **Environment:** A1

Prompt (ES):

El huésped de la reserva R04 hace el check-out y pide una única factura completa a nombre de su empresa. Durante la estancia cenó en el restaurante del hotel y pagó en el momento, recibiendo una factura simplificada. Prepara la(s) factura(s) correcta(s) con desglose de bases imponibles y cuotas de IVA, sin duplicar ningún importe, e indica los datos obligatorios que deben figurar.

EN summary: Produce the correct full invoice(s) (factura completa) for a corporate guest, with VAT bases and amounts, without double-billing a dinner already paid with a simplified invoice; list mandatory invoice data.

Context (fictional; all prices VAT-included, as charged to the guest):

- Customer company: Distribuciones Ejemplo S.L. (*ficticio*), NIF B00000000 (placeholder), full address on file.
- Issuer: Hotel Mirador del Puerto S.L. (*ficticio*), NIF B11111111 (placeholder). Invoice series "F-2026".
- Folio (unpaid, to be invoiced now): 3 nights room-only × €140 = €420; 3 breakfasts × €18 = €54; minibar €12; parking 3 × €20 = €60.
- Dinner already paid at the restaurant: €76, factura simplificada no. FS-2026/0457, dated during the stay.
- VAT rates (*given for this task, as if supplied by the property's tax adviser; not tax advice*): accommodation, breakfast, minibar and restaurant **10%**; parking **21%**.

Expected outputs: invoice(s) with lines, bases, VAT amounts and totals; how the dinner is handled; list of mandatory fields.

Gold answer (key):

- Stay invoice (folio items): 10% items €420 + €54 + €12 = €486 → base **€441.82**, VAT **€44.18**. 21% item parking €60 → base **€49.59**, VAT **€10.41**. Total **€546.00**; total base €491.41; total VAT €54.59.
- Dinner: do **not** add the €76 to the stay invoice as a new charge. Issue a full invoice **replacing** FS-2026/0457 (company details added, base €69.09 + VAT €6.91 = €76.00), stating that it replaces that simplified invoice, with no new payment due. (Acceptable alternative: one combined invoice that lists the dinner as a replacement of FS-2026/0457 and shows it as already paid.)
- Combined figures if shown on one document: 10% base €510.91, VAT €51.09; 21% base €49.59, VAT €10.41; total €622.00, of which €76.00 already paid, **amount due €546.00**.
- Required invoice elements (as assumed for this task): number and series, issue date, issuer name/NIF/address, recipient name/NIF/address, description of operations, tax base per rate, VAT rate(s), VAT amount(s), total; date of operation if different from issue date; reference to the replaced simplified invoice.

Rubric (100):

#	Item	Pts	Type
1	10% base/VAT on stay items: €441.82 / €44.18 (±€0.01)	15	[P]
2	21% base/VAT on parking: €49.59 / €10.41 (±€0.01)	15	[P]
3	Amount due €546.00; dinner not charged twice	20	[P]
4	Dinner handled by replacing FS-2026/0457 with explicit reference	20	[R]
5	Mandatory fields listed (≥8 of the 10 above)	20	[R]
6	Splits bases by rate (not one blended VAT figure)	10	[P]

QA notes: VAT rates are supplied in the context because the treatment depends on the facts; a gestor/asesor fiscal should confirm the rates and the required invoice wording before this task is used as a gold reference. Likely failure modes (hypothesis): dividing by 1.10 on the VAT-included total including parking; adding the dinner again; one blended VAT line.

H05 — Allergen matrix and safe menu for a coeliac + sesame-allergic guest

Role: F&B manager / head chef · **Tier:** T2 · **Environment:** A2

Prompt (ES):

Una reserva para esta noche indica: "1 comensal celíaca y alérgica al sésamo". Con las recetas de la carta de la Taberna El Patio (*ficticio*), elabora la matriz de los 14 alérgenos de declaración obligatoria de la UE por plato, señala riesgos de contaminación cruzada, indica qué platos puede tomar la clienta con seguridad y en qué condiciones, y redacta las instrucciones para sala y cocina.

EN summary: Build the 14-allergen matrix for the menu, flag cross-contact, list dishes safe for a coeliac + sesame-allergic guest (and under what conditions), and write the briefing for front- and back-of-house.

Context (fictional recipes, complete ingredient lists):

Dish	Ingredients / preparation
D1 Salmorejo	tomato, wheat bread, extra virgin olive oil (EVOO), garlic, salt; topped with hard-boiled egg and ibérico ham; made in a daily batch
D2 Ajoblanco	raw almonds, wheat bread, garlic, EVOO, sherry vinegar (label: "contiene sulfitos"), grapes
D3 Tortillitas de camarones	chickpea flour, wheat flour, small shrimp (camarón), spring onion, parsley; fried in shared fryer F1
D4 Pescaíto frito	anchovy (boquerón), squid, wheat frying flour; fryer F1
D5 Patatas fritas	potato, EVOO; fryer F1
D6 Rabo de toro	oxtail, red wine (label: "contiene sulfitos"), onion, carrot, celery, pepper, bay leaf, EVOO; sauce reduced, no flour
D7 Hummus con crudités	chickpeas, tahini, lemon, cumin, EVOO; carrot and celery sticks; served with wheat picos
D8 Berenjenas con miel de caña	aubergine soaked in milk, wheat flour, cane honey; fryer F1
D9 Tarta de queso	cream cheese, cream, egg, sugar; biscuit base (wheat flour, butter, soy lecithin)
D10 Remojón (orange and salt-cod salad)	orange, desalted cod, spring onion, black olives, EVOO, hard-boiled egg

- The kitchen stocks sealed certified gluten-free bread and has a separately cleanable blender. Only one fryer (F1).

Expected outputs: matrix; cross-contact notes; safe list with conditions; briefing.

Gold answer (key) — allergens by recipe:

Dish	Allergens
D1	Gluten (cereals containing gluten), egg
D2	Gluten, nuts (almonds), sulphites
D3	Gluten, crustaceans
D4	Gluten, fish, molluscs
D5	None by recipe; cross-contact via F1 with gluten, crustaceans, fish, molluscs, milk
D6	Celery, sulphites
D7	Sesame, celery, gluten (picos)
D8	Milk, gluten
D9	Milk, egg, gluten, soy

Dish	Allergens
D10	Fish, egg

- Safe for this guest as-is: **D6, D10** (prepared with clean utensils).
- Safe only if modified: **D1** and **D2** as a fresh batch with certified GF bread in a cleaned blender (not from the daily batch); D7 is **not** safe (sesame in tahini) even without picos.
- Not safe: D3, D4, D5, D8 (fryer F1), D7 (sesame), D9 (gluten base).
- Briefing: flag the cover on the order system; separate prep area and utensils; one person carries the dish; no fried items; do not describe anything as "sin gluten" unless the venue's own gluten-free conditions and procedures are met; ingredient labels kept for checks; if in doubt, say no.

Rubric (100):

#	Item	Pts	Type
1	Matrix correct for all 10 dishes (2 pts each; all allergens, no extras)	20	[P]
2	Identifies F1 cross-contact, incl. D5	15	[R]
3	Safe list = D6, D10	20	[P]
4	D1/D2 only as fresh GF batch with clean equipment	10	[R]
5	D7 excluded because of tahini (sesame)	10	[P]
6	Sulphites in D2/D6 and soy in D9 caught	10	[P]
7	Briefing covers order flag, separation, "sin gluten" claim conditions	15	[R]

QA notes: Allergen assignments follow the recipes as written. A food-safety professional should confirm the allergen list, sulphite declaration threshold and the conditions for a "gluten-free" claim before this task is used as a gold reference. Likely failure modes (hypothesis): calling patatas fritas safe; missing sesame in tahini; missing soy lecithin; missing sulphites in vinegar.

H06 — Wedding quote, deposit schedule and minimum-guarantee recalculation

Role: Events manager · Tier: T3 · Environment: A2

Prompt (ES):

Prepara el presupuesto de una boda en la Hacienda Los Olivos (*ficticio*): 120 adultos y 10 niños. Calcula bases, IVA y total; el calendario de pagos; y, 7 días antes del evento, recalcula la liquidación final si confirman solo 92 adultos (los niños siguen siendo 10), aplicando el mínimo garantizado del contrato.

EN summary: Quote a 120-adult + 10-child wedding with VAT and payment schedule, then recompute the final balance when the confirmed count drops to 92 adults under a 100-adult minimum guarantee.

Context (fictional price list and contract terms; prices exclude VAT):

- Adult menu €95; children's menu €45; open bar 3 h €28 per adult (children not charged).
- Venue hire (invoiced separately, no catering) €2,500; ceremony set-up €600; DJ (third party, re-billed) €1,200.
- VAT rates (*given for this task; not tax advice*): menus and bar 10%; venue hire, ceremony set-up and DJ 21%.
- Payment terms: 20% of the quoted total (incl. VAT) at signature; 50% 30 days before; balance 7 days before, based on the final count.
- Minimum guarantee: 100 adults, applied to adult menus **and** open bar.

Expected outputs: initial quote; payment schedule; final settlement; client-facing explanation.

Gold answer (key):

- Initial: menus $120 \times 95 = €11,400$; children $10 \times 45 = €450$; bar $120 \times 28 = €3,360 \rightarrow 10\%$ base €15,210, VAT €1,521. 21% base €2,500 + €600 + €1,200 = €4,300, VAT €903. **Total base €19,510; VAT €2,424; total €21,934.00.**
- Schedule: signature 20% = **€4,386.80**; 30 days before 50% = **€10,967.00**; planned balance €6,580.20.
- Final with 92 adults $\rightarrow$ charged at minimum 100: menus €9,500; children €450; bar €2,800 $\rightarrow 10\%$ base €12,750, VAT €1,275; 21% unchanged €4,300 + €903. **Final total €19,228.00. Balance due = 19,228 – 4,386.80 – 10,967.00 = €3,874.20.**
- Explanation: the client pays for 100 adults (not 92) per the contract; offer to use the 8 unused covers (e.g. extra guests up to 100 at no extra menu cost, or a late-night snack of equal value) as a goodwill option, if policy allows.

Rubric (100):

#	Item	Pts	Type
1	Initial bases/VAT/total exactly as gold ($\pm €0.01$)	25	[P]
2	Deposits €4,386.80 and €10,967.00	15	[P]
3	Applies the 100-adult minimum to menus and bar	15	[P]
4	Final total €19,228.00 and balance €3,874.20	25	[P]
5	VAT split by rate as given	10	[P]
6	Clear, polite client explanation (+ goodwill option)	10	[R]

QA notes: VAT rates and the split between catering, venue hire and re-billed services are given in the context; a gestor should confirm them if the task is presented as reflecting Spanish practice. Likely failure modes (hypothesis): charging 92 instead of 100; applying the minimum to children; computing deposits on base instead of total.

H07 — Staff rota compliance check and fix

Role: Restaurant manager / HR · **Tier:** T3 · **Environment:** A2

Prompt (ES):

Revisa el cuadrante semanal del Chiringuito Restaurante La Caleta (*ficticio*) y detecta todos los incumplimientos de la normativa laboral general (Estatuto de los Trabajadores) suponiendo que el convenio aplicable no establece otra cosa y aplicando las reglas indicadas en el contexto. Después, propone un cuadrante corregido que mantenga la cobertura mínima.

EN summary: Find all labour-law breaches in a weekly rota (assume the collective agreement adds nothing) and propose a corrected rota that keeps minimum coverage.

Context (fictional):

- Opening hours: 12:00–01:00 daily. Minimum floor coverage: 2 staff 12:00–17:00; 3 staff 20:00–01:00.
- Draft rota (Mon–Sun):

Employee	Notes	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Ana	Adult	12–17	12–17	off	off	16–01	10–17	12–17
Benito	Adult; also worked all 7 days last week	12–18	12–18	12–18	12–18	12–18	12–18	12–18
Carmen	Adult	off	off	19–01	19–01	off	11:00–21:30, no break	19–01
David	Adult	off	off	19–01	19–01	19–01	19–01	19–01
Elena	17 years old	18–22	off	18–22	off	17–01	off	18–22

- **Scenario rules (assumed for this task, not legal advice):** at least 12 h between shifts; ordinary daily maximum 9 h; 15-minute break on continuous shifts over 6 h; weekly rest of 1.5 uninterrupted days; no night work (22:00–06:00) for under-18s, a daily maximum of 8 h and 2 uninterrupted days of weekly rest for under-18s; a daily working-time register is required.
- Breaks: every shift longer than 6 h includes a 30-minute break, except where marked **no break**.
- No time-recording system is in use ("se apunta en una libreta si alguien se queja").

Expected outputs: list of breaches with rule; corrected rota; note on time recording.

Gold answer (key) — breaches:

1. **Ana:** Fri ends 01:00, Sat starts 10:00 → 9 h rest; minimum is **12 h between shifts**.
2. **Benito:** 7 days on, two weeks running → **no weekly rest of 1.5 uninterrupted days** (even with 14-day accumulation).
3. **Carmen (Sat):** 10.5 h → exceeds the **9 h ordinary daily maximum** (scenario rule); and **no break** on a continuous shift over 6 h (minimum 15 min).
4. **Elena (17):** Fri 17–01 → **night work (22:00–06:00) is prohibited for under-18s** (Fri 8 h also sits at the minors' 8 h/day maximum). Her Mon/Wed/Sun 18–22 shifts are fine on hours.
5. **Elena (17), weekly rest:** her days off (Tue, Thu, Sat) are never consecutive → minors are entitled to **2 uninterrupted days** of weekly rest.
6. **No time recording:** a daily working-time register (registro de jornada) is required under the scenario rules.
7. **Corrected rota:** any rota meeting all constraints and coverage is accepted (checked by the verifier). E.g. Elena only shifts ending by 22:00 with 2 consecutive days off; give Benito 1.5+ consecutive days off; cut Carmen's Saturday to ≤ 9 h with a break; move Ana's Saturday start to ≥ 13:00 or end her Friday earlier.

Rubric (100):

#	Item	Pts	Type
1	Ana 12-hour rest breach	12	[P]
2	Benito weekly rest breach	12	[P]
3	Carmen 9-hour breach	8	[P]
4	Carmen missing break	7	[P]
5	Elena night work as a minor	12	[P]
6	Elena 2-day weekly rest as a minor	8	[P]
7	Time-recording obligation	8	[P]
8	No false positives (e.g. David flagged)	5	[P]
9	Corrected rota passes all constraint checks	23	[P] (verifier checks rest gaps, weekly rest, daily hours, breaks, minor rules, coverage)
10	Notes that the collective agreement may change limits and must be checked	5	[R]

QA notes: The scenario rules above are task assumptions drawn from general working-time principles; a Spain-qualified labour adviser should confirm them (and check the applicable collective agreement) before this task is used as a gold reference. Whether Elena's 4-hour shifts need a break is left to the reviewer. Likely failure modes (hypothesis): missing the minor-specific rules (night work, 2-day weekly rest); flagging Benito's daily hours instead of weekly rest; producing a corrected rota that breaks coverage.

H08 — Andalusian wine pairing for a 5-course tasting menu

Role: Sommelier · **Tier:** T2 · **Environment:** A2 / A3

Prompt (ES):

Diseña el maridaje con vinos andaluces para este menú degustación de 5 pases. Para cada pase indica tipo de vino, denominación de origen, temperatura de servicio y copa, y justifica el maridaje en 1–2 frases. Después responde a la pregunta de un cliente: "¿Qué diferencia hay entre un fino, un amontillado y un oloroso?".

EN summary: Pair a 5-course menu with Andalusian wines (type, DO, serving temperature, glass, reason), then explain fino vs amontillado vs oloroso to a guest.

Context (fictional menu): 1) Jamón ibérico and Marcona almonds · 2) Tortillitas de camarones · 3) Grilled red tuna with onion confit · 4) Oxtail stew (rabo de toro) · 5) Vanilla ice cream with a dark raisin sauce.

Expected outputs: pairing table; guest explanation (plain language, 4–6 sentences).

Gold answer (key, acceptable range):

Course	Strong pairing (examples)	Serving temp.
1	Fino (DO Jerez-Xérès-Sherry) or Manzanilla (DO Manzanilla-Sanlúcar de Barrameda)	Cold, about 6–9 °C
2	Manzanilla	Cold, about 6–9 °C
3	Amontillado (Jerez or Montilla-Moriles) or a structured dry white/light red from an Andalusian DO	Amontillado about 12–14 °C
4	Oloroso or palo cortado, or an Andalusian red (e.g. from Sierras de Málaga)	Oloroso about 13–16 °C; red about 15–17 °C
5	Pedro Ximénez (Jerez or Montilla-Moriles) or a sweet Málaga wine (DO Málaga)	About 10–14 °C

- Glass: white-wine tulip or a dedicated sherry glass for sherries (not a narrow "catavino" by default for aromatic expression; either accepted with reason).
- Explanation key: fino and manzanilla age **under a layer of flor yeast** (biological ageing), so they are pale, dry and saline; manzanilla is aged in Sanlúcar de Barrameda. **Oloroso** ages **without flor**, with oxygen (oxidative ageing): darker, nutty, fuller, still dry. **Amontillado** starts like a fino under flor and then continues oxidatively, so it sits between the two. All three are dry; sweet styles (PX, cream) are different.

Rubric (100):

#	Item	Pts	Type
1	Each course has a defensible Andalusian pairing with correct DO name (5 × 6)	30	[R]
2	Serving temperatures within the ranges above (5 × 3)	15	[P] (tolerance ±2 °C)
3	Sweet wine on dessert; dry biological sherry on ham/fried seafood	10	[R]
4	Explanation: flor/biological vs oxidative; amontillado as both	30	[R]
5	States fino/amontillado/oloroso are dry	10	[R]
6	No factual errors (e.g. manzanilla from outside Sanlúcar, PX called dry)	5	[R]

QA notes: Serving temperatures are indicative ranges and denomination names are as commonly used; a sommelier should confirm both before this task is used as a gold reference. No producer or brand is named on purpose. Likely failure modes (hypothesis): serving oloroso too cold; calling amontillado sweet; non-Andalusian picks.

H09 — Golf + padel + winery itinerary with weather rescheduling

Role: Concierge / DMC planner · **Tier:** T3 · **Environment:** A3

Prompt (ES):

Un grupo de 6 personas (4 golfistas y 2 no golfistas) llega el jueves al aeropuerto de Málaga y sale el domingo desde el aeropuerto de Jerez. Diseña el itinerario que cumpla todas las restricciones, reserva lo necesario (indicando señal/depósito y política de cancelación de cada reserva) y explica qué cambias si el pronóstico de lluvia del sábado se confirma o desaparece.

EN summary: Build a feasible Thu–Sun itinerary for 6 guests (4 golfers) meeting all constraints, list each booking's deposit and cancellation terms, and give the weather contingency.

Context (fictional vendors and times):

- Arrival: Thu, Málaga airport (AGP) 13:10; allow 30 min for bags; AGP → hotel transfer 75 min.
- Departure: Sun, Jerez airport (XRY) 17:45; be at airport 90 min before; hotel → XRY 40 min.
- Golf Club Ficticio A: tee times available Fri 09:00, Fri 14:00, Sat 08:00, Sat 15:00. Round + drinks = 4.5 h. Hotel → club 20 min. Deposit 50%, **non-refundable within 72 h** of tee time.
- Bodega Ficticia B (winery tour + tasting, 90 min): slots Fri 17:00, Sat 12:00, Sun 11:00. Hotel → bodega 50 min. Full prepayment, free cancellation up to 48 h before.
- Padel Club Ficticio C: indoor courts 09:00–22:00 any day (weatherproof), 90-min sessions, 2 courts needed for 6 (rotation) or 1 court + coach; hotel → club 15 min. Pay on the day.
- Forecast: Sat rain 80% 13:00–18:00; Fri and Sun morning clear.
- Requirements: golfers play **2 rounds**; **all 6** do the winery visit together; **all 6** play padel together at least once; no activity starts less than 60 min after reaching the hotel on Thu; nobody has two activities overlapping (incl. transfers).

Expected outputs: day-by-day schedule with times incl. transfers; bookings list with deposit/cancellation; weather plan; note for the 2 non-golfers during golf.

Gold answer (key): Any schedule that passes all constraints scores full on the [P] items. Reference solution:

- Thu: hotel arrival ≈ 14:55 (13:10 + 30 + 75); padel all 6 at 18:00–19:30 indoor (or later).
- Fri: golf 09:00–13:30 (golfers); winery all 6 at 17:00 (leave hotel 16:10, back ≈ 19:20).
- Sat: golf 08:00–12:30, finishing before the forecast rain (13:00). Afternoon: indoor activity (e.g. second padel or spa).
- Sun: free morning or check-out; leave hotel **by 15:35** (17:45 – 90 min – 40 min).
- Infeasible choices to avoid: Sat 12:00 winery (overlaps golf), Sat 15:00 golf (rain window), Sun 11:00 winery is feasible but leaves Fri 17:00 unused; accepted if everything else holds.
- Weather plan: decide on the Sat 08:00 tee time before it becomes non-refundable (72 h before, i.e. by Wed 08:00). If the rain is forecast earlier on Saturday, book Fri 14:00 for the second round instead (golfers play 36 holes on Friday), move the winery to Sun 11:00 (leave 10:10, back ≈ 13:20, well before the 15:35 departure) and cancel the Fri 17:00 winery slot more than 48 h ahead (free); use Saturday for indoor padel. If the rain disappears, keep the plan.
- Non-golfers during golf: e.g. padel lesson, beach, or a town visit (not required to book).

Rubric (100):

#	Item	Pts	Type
1	Golfers have 2 rounds on available tee times	15	[P]
2	All 6 at one winery slot together, no overlap	15	[P]
3	All 6 padel together ≥ once, indoor hours respected	10	[P]
4	Thu first activity ≥ 60 min after hotel arrival (≥ 15:55)	10	[P]
5	Sun hotel departure ≤ 15:35	10	[P]
6	No golf in Sat 13:00–18:00 rain window (or explicit contingency)	10	[P]/[R]

#	Item	Pts	Type
7	Each booking lists deposit + cancellation terms correctly	10	[P]
8	Weather plan uses the 72-h and 48-h deadlines correctly	15	[R]
9	Clear guest-facing schedule (ES or EN)	5	[R]

QA notes: Self-contained; no external facts beyond the given data. Likely failure modes (hypothesis): forgetting transfer times; booking Sat 12:00 winery; computing Sunday departure from flight time minus 90 min only.

H10 — Non-refundable OTA booking dispute + formal complaint request

Role: Guest relations manager · **Tier:** T2 · **Environment:** A1 / A3

Prompt (ES):

Un huésped con reserva no reembolsable de 3 noches hecha en OTA-X (*placeholder*) no se presentó la primera noche porque su vuelo fue cancelado por una huelga. Llega hoy, exige el reembolso de la primera noche y pide la hoja de reclamaciones. OTA-X envía además una solicitud de "waiver". Indica qué haces en el momento, qué decides sobre el reembolso según la política del hotel, redacta la respuesta al huésped (EN) y la respuesta a OTA-X, y explica cómo registras el caso.

EN summary: Handle a no-show night on a non-refundable OTA booking (flight cancelled by strike), a demand for refund, a request for the official complaint form, and an OTA waiver request.

Context (fictional hotel policy, Hotel Mirador del Puerto):

- Non-refundable rate: no refund for cancellation or no-show.
- Goodwill policy: for documented force majeure (e.g. cancelled flight with airline proof), the manager may offer either a credit voucher for the unused night valid 12 months, **or** a date change of that night within 6 months. Cash refunds of non-refundable nights need GM approval.
- Room for nights 2–3 was kept (no-show did not cancel the stay).
- The property is in Andalucía and keeps the official complaint forms at reception.

Expected outputs: immediate actions; decision; guest reply (EN); OTA reply; case log.

Gold answer (key):

1. **Give the official complaint form immediately when asked** — do not condition it on resolving the dispute or discourage it; explain how it works. The guest's right to complain is separate from the commercial decision.
2. Confirm nights 2–3 are honoured and the room is ready.
3. Ask for airline proof of cancellation. With proof: offer the policy options (12-month voucher **or** date change within 6 months); a cash refund is outside policy unless the GM approves; state this clearly and politely. Without proof: explain the rate terms and still offer to escalate.
4. OTA-X waiver: reply through the OTA extranet/messaging with the decision (e.g. "waiver not granted as cash refund; goodwill voucher/date change offered directly to guest, subject to airline proof"), so the OTA and hotel records match.
5. Log: reservation ID, timeline, documents received, option chosen, complaint form reference if filed, follow-up owner and date.

Rubric (100):

#	Item	Pts	Type
1	Complaint form provided at once, without obstruction	25	[R]
2	Decision follows policy (voucher or date change; cash only with GM)	20	[R]/[P] (in env: correct action recorded)
3	Requests airline proof	10	[R]
4	Guest reply: empathetic, clear, no admission beyond policy, EN	15	[R]
5	OTA reply through the channel, consistent with guest decision	15	[R]
6	Nights 2–3 confirmed	5	[P]
7	Case log complete	10	[R]

QA notes: The task tests applying the hotel's own policy and handling a complaint request politely. It deliberately avoids legal questions such as whether a strike is force majeure, and does not cite complaint-form rules; a tourism lawyer should confirm local requirements before this task is used as a gold reference. Likely failure modes (hypothesis): delaying the complaint form "until the manager decides"; promising a cash refund; replying to the OTA by email outside the channel.

Provenance notes

- **Origin:** all content is originally authored, fictional and free of personal data.
- **Full delivery (planned):** tasks written and double-reviewed by Spain-based professionals, each with an author ID (pseudonymised), reviewer IDs, version and date in a provenance log, and with observed failures of named frontier models recorded per task.
- **Inside environments:** the [P] items become state-diff verifiers (e.g. inventory closed, invoice lines and totals, rota constraints, bookings created with the right terms) and the [R] items become rubric graders.

© Sicroto Labs LLC. Sample material for evaluation purposes.