

A hundred questions this archive can answer

Prepared 7 September 2026 by the Lunenburg Budget Project, an independent archive of the town and school budgets at lunenburgbudgetproject.org. Every figure quoted here is traceable to a document published on that site.

Generated by `scripts/build_question_bank.py`. **Do not edit.**

Every query below is run against the database on each build, and the build fails if one errors or returns nothing. 107 questions across 10 subjects.

Run any of them yourself:

```
`` curl -s -X POST https://lunenburgbudgetproject.org/api/query \ -H 'content-type: application/json' \ -d '{"sql": "SELECT fy, COUNT(*) FROM v_staff_roster WHERE role_category = 'paraprofessional' GROUP BY fy"}' ``
```

Or download the database: <https://lunenburgbudgetproject.org/data/lunenburg.db>. Read <https://lunenburgbudgetproject.org/api/schema> first — it states the grain of every table and the specific ways to get a confident wrong answer out of this data.

These are not answers. The numbers move when the data does, so this shows the shape of each result — its columns and a row or two — rather than repeating figures that would go stale. Several entries exist to demonstrate a rule rather than to be interesting; those carry a note and are the ones worth copying.

The school budget

What did the district budget in total, in each year and at each stage?

```
`` sql SELECT fy, stage, ROUND(SUM(value)) AS total FROM budget_figure GROUP BY fy, stage ORDER BY fy, stage ``
```

Returns `fy`, `stage`, `total` — for example: `fy=2014, stage=actual, total=16500031.0`

A STAGE is not a period. `proposed`, `settled` and `actual` are three different documents about the same year, and mixing them is the error rule 1 exists for.

Which budget lines grew fastest across the years the archive holds?

```
`` sql WITH span AS (SELECT line_key, MIN(fy) AS first_fy, MAX(fy) AS last_fy FROM budget_figure WHERE stage='settled' GROUP BY line_key HAVING last_fy > first_fy) SELECT b.label, s.first_fy, s.last_fy, ROUND(a.value) AS started, ROUND(z.value) AS ended, ROUND(100.0*(z.value-a.value)/a.value,1) AS pct FROM span s JOIN budget_line b USING (line_key) JOIN budget_figure a ON a.line_key=s.line_key AND a.fy=s.first_fy AND a.stage='settled' JOIN budget_figure z ON z.line_key=s.line_key AND z.fy=s.last_fy AND z.stage='settled' WHERE a.value > 1000 ORDER BY pct DESC LIMIT 20 ``
```

Returns `label`, `first_fy`, `last_fy`, `started`, `ended`, `pct` — for example: `label=Business Office/Clerical`, `first_fy=2020`, `last_fy=2026`, `started=7952.0`, `ended=106500.0`, `pct=1239.3`

Both ends are the SAME stage. A rate measured from an actual to a budget is partly growth and partly the step between them, which is rule 1.

Which lines do two documents state differently, and by how much?

```
``sql SELECT label, fy, stage, ROUND(spread) AS spread FROM v_budget_disagreement WHERE spread > 0 ORDER BY spread DESC LIMIT 20 ``
```

Returns `label`, `fy`, `stage`, `spread` — for example: `label=Special Ed Tuitions/Private`, `fy=2025`, `stage=proposed`, `spread=434173.0`

The documents disagree with themselves by up to 1.5%, which is larger than most variances anybody wants to measure.

What does each FY27 scenario total?

```
``sql SELECT variant, ROUND(SUM(value)) AS total, COUNT(*) AS lines FROM budget_figure WHERE fy=2027 AND variant IS NOT NULL GROUP BY variant ORDER BY total DESC ``
```

Returns `variant`, `total`, `lines` — for example: `variant=Restoration`, `total=28364451.0`, `lines=250`

How many budget lines are there in each section of the budget?

```
``sql SELECT section, COUNT(*) AS lines FROM budget_line GROUP BY section ORDER BY lines DESC ``
```

Returns `section`, `lines` — for example: `section=None`, `lines=350`

Which function groups hold the most budget lines?

```
``sql SELECT function_group, COUNT(*) AS lines FROM budget_line WHERE function_group <> '' GROUP BY function_group ORDER BY lines DESC LIMIT 15 ``
```

Returns `function_group`, `lines` — for example: `function_group=2415 - H.S. Other Instr. Materials`, `lines=16`

What is in the FY27 workbook, by column?

```
``sql SELECT column_kind, COUNT(*) AS rows, ROUND(SUM(value)) AS total FROM workbook_figure WHERE row_kind='line' GROUP BY column_kind ORDER BY rows DESC ``
```

Returns `column_kind`, `rows`, `total` — for example: `column_kind=actual`, `rows=748`, `total=69737410.0`

Filter `row_kind='line'`: the sheet's own TOTAL rows are loaded too, and summing without the filter double-counts roughly fourfold.

Which budget lines appear in the workbook but not in the line catalogue?

```
``sql SELECT DISTINCT w.line_key FROM workbook_figure w LEFT JOIN budget_line b USING (line_key) WHERE b.line_key IS NULL LIMIT 20 ``
```

Returns `line_key` — for example: `line_key=total` actuals & budget:

How many figures rest on a document that two sources report differently?

```
``sql SELECT fy, COUNT(*) AS figures FROM budget_figure WHERE documents_disagree=1 GROUP BY fy ORDER BY fy ``
```

Returns `fy`, `figures` — for example: `fy=2014`, `figures=2`

What is the total salary line in each year, and does the stage change it?

```
``sql SELECT fy, stage, total FROM total_salaries_history ORDER BY fy, stage ``
```

Returns `fy`, `stage`, `total` — for example: `fy=2014`, `stage=actual`, `total=11044481`

And total expenses?

```
``sql SELECT fy, stage, total FROM total_expenses_history ORDER BY fy, stage ``
```

Returns `fy`, `stage`, `total` — for example: `fy=2014`, `stage=actual`, `total=5146641`

What is the biggest single line in the budget, in each year?

```
``sql SELECT f.fy, b.label, ROUND(MAX(f.value)) AS value FROM budget_figure f JOIN budget_line b USING (line_key) WHERE f.stage='settled' GROUP BY f.fy ORDER BY f.fy DESC ``
```

Returns `fy`, `label`, `value` — for example: `fy=2026`, `label=Health Insurance`, `value=3701195.0`

How many lines does each budget document state?

```
``sql SELECT doc_id, COUNT(*) AS figures, COUNT(DISTINCT fy) AS years FROM budget_figure GROUP BY doc_id ORDER BY figures DESC LIMIT 15 ``
```

Returns `doc_id`, `figures`, `years` — for example: `doc_id=sources/district-budget/text/budget-hearing-fy20-proposed-lps-budget.txt`, `figures=1148`, `years=6`

Which lines exist in one scenario but not another?

```
``sql SELECT line_key, COUNT(DISTINCT variant) AS variants FROM budget_figure WHERE fy=2027 AND variant IS NOT NULL GROUP BY line_key ORDER BY variants LIMIT 15 ``
```

Returns `line_key`, `variants` — for example: `line_key=hs audio visual supplies`, `variants=1`

What does the FY27 workbook say each line was in FY25?

```
``sql SELECT line_key, column_kind, ROUND(value) AS value FROM workbook_figure WHERE fy=2025 AND row_kind='line' ORDER BY value DESC LIMIT 20 ``
```

Returns `line_key`, `column_kind`, `value` — for example: `line_key=health insurance`, `column_kind=actual`, `value=3248744.0`

Special education

How many special education paraprofessionals were budgeted, by school and year?

```
``sql SELECT fy, stage, ps, es, ms, hs, total FROM sped_para_history ORDER BY fy, stage ``
```

Returns `fy, stage, ps, es, ms, hs, total` — for example: `fy=2014, stage=actual, ps=153425, es=0, ms=162721, hs=50819, total=366965`

This is the line the 12.8% escalator rests on, and it is dollars, not people.

And special education teachers?

```
``sql SELECT fy, stage, ps, es, ms, hs, total FROM sped_teacher_history ORDER BY fy, stage ``
```

Returns `fy, stage, ps, es, ms, hs, total` — for example: `fy=2017, stage=actual, ps=473861, es=307620, ms=375364, hs=380738, total=1537583`

What has out-of-district tuition done, year by year?

```
``sql SELECT fy, stage, private, collaborative, total FROM ood_tuition_history ORDER BY fy, stage ``
```

Returns `fy, stage, private, collaborative, total` — for example: `fy=2014, stage=actual, private=1025404, collaborative=236285, total=1261689`

How many children were placed outside the district, and where?

```
``sql SELECT fy, as_of, total, collaborative, day, residential FROM placement_counts ORDER BY fy ``
```

Returns `fy, as_of, total, collaborative, day, residential` — for example: `fy=2011, as_of=2011-03-01, total=15, collaborative=, day=13, residential=2`

A count of children placed. It says nothing about which fund paid or what a placement cost, so it does not settle the money.

Do the placement counts tie to their own parts, and to the prior year?

```
``sql SELECT fy, parts_tie, chain_agrees, report_says_prior_year FROM placement_counts ORDER BY fy ``
```

Returns `fy, parts_tie, chain_agrees, report_says_prior_year` — for example: `fy=2011, parts_tie=yes, chain_agrees=n/a, report_says_prior_year=`

What has special education transportation cost, by year?

```
``sql SELECT fy, stage, system, total FROM sped_transport_history ORDER BY fy, stage ``
```

Returns `fy, stage, system, total` — for example: `fy=2015, stage=actual, system=480536, total=480536`

Does each year of the placement series agree with what the next report says of it?

```
``sql SELECT fy, total, report_says_prior_year, chain_agrees FROM placement_counts ORDER BY fy
``
```

Returns `fy`, `total`, `report_says_prior_year`, `chain_agrees` — for example: `fy=2011, total=15, report_says_prior_year=, chain_agrees=n/a`

Two checks travel with this series: the parts sum to the total, and each year states the previous year's figure. `n/a` is the first year, which has nothing before it.

The town's books

What did each department spend against its budget, in the latest period held?

```
``sql SELECT a.dept, a.name, l.fy, l.period, ROUND(l.revised) AS revised, ROUND(l.expended) AS
expended, ROUND(l.available) AS available FROM ledger_snapshot l JOIN account a USING
(account_id) WHERE a.level='department' ORDER BY l.fy DESC, l.period DESC, l.revised DESC LIMIT
20 ``
```

Returns `dept`, `name`, `fy`, `period`, `revised`, `expended`, `available` — for example: `dept=300, name=SCHOOL DEPARTMENT, fy=2026, period=9, revised=26323868.0, expended=15736641.0, available=8919184.0`

Which departments are spending faster than the year is elapsing?

```
``sql SELECT dept, name, fy, period, ROUND(year_elapsed,2) AS year_elapsed, ROUND(spent_share,2)
AS spent_share, ROUND(pace_gap,2) AS pace_gap FROM v_burn WHERE pace_gap IS NOT NULL ORDER BY
pace_gap DESC LIMIT 20 ``
```

Returns `dept`, `name`, `fy`, `period`, `year_elapsed`, `spent_share`, `pace_gap` — for example: `dept=None, name=PS TUITION, fy=2026, period=9, year_elapsed=0.75, spent_share=7.53, pace_gap=6.78`

What funds does the town keep, and what restricts them?

```
``sql SELECT kind, COUNT(*) AS funds FROM fund GROUP BY kind ORDER BY funds DESC ``
```

Returns `kind`, `funds` — for example: `kind=None, funds=59`

What moved through the special revenue funds in each year?

```
``sql SELECT fund, fy, period, ROUND(opening_balance) AS opening, ROUND(revenue) AS revenue,
ROUND(expenditure) AS spent, ROUND(closing_balance) AS closing FROM fund_activity ORDER BY fy
DESC, revenue DESC LIMIT 20 ``
```

Returns `fund`, `fy`, `period`, `opening`, `revenue`, `spent`, `closing` — for example: `fund=2200, fy=2026, period=9, opening=None, revenue=572231.0, spent=521910.0, closing=287771.0`

How many accounts are there at each level of the chart?

```
``sql SELECT level, account_type, COUNT(*) AS accounts FROM account GROUP BY level, account_type ORDER BY accounts DESC ``
```

Returns `level`, `account_type`, `accounts` — for example: `level=account`, `account_type=expense`, `accounts=692`

Which grants did the district receive, and who owns them?

```
``sql SELECT fy, kind, name, ROUND(amount) AS amount, owner FROM grant_award ORDER BY fy DESC, amount DESC LIMIT 20 ``
```

Returns `fy`, `kind`, `name`, `amount`, `owner` — for example: `fy=FY21-24`, `kind=federal`, `name=ESSER 3`, `amount=1351034.0`, `owner=`

What does the ledger hold for each fiscal year and period?

```
``sql SELECT fy, period, COUNT(*) AS rows, COUNT(DISTINCT account_id) AS accounts FROM ledger_snapshot GROUP BY fy, period ORDER BY fy, period ``
```

Returns `fy`, `period`, `rows`, `accounts` — for example: `fy=2026`, `period=9`, `rows=348`, `accounts=346`

Period 13 is the year-end close, after purchase orders are cleared. Period 12 is not the end of the year.

Which function codes can be compared between the budget and the ledger?

```
``sql SELECT function_code, fy, period, ROUND(ledger_revised) AS revised, ROUND(ledger_expended) AS expended, budget_lines FROM v_function_budget_vs_ledger ORDER BY ledger_revised DESC LIMIT 20 ``
```

Returns `function_code`, `fy`, `period`, `revised`, `expended`, `budget_lines` — for example: `function_code=2305`, `fy=2026`, `period=12`, `revised=7929717.0`, `expended=7867027.0`, `budget_lines=7`

This is the level at which the two systems join. Below it they do not: the town shortens account names to ten characters.

How much did each function group budget and spend across all years?

```
``sql SELECT function_group, years, ROUND(budgeted) AS budgeted, ROUND(spent) AS spent, ROUND(net) AS net, worst_year FROM variance_by_group ORDER BY ABS(net) DESC LIMIT 20 ``
```

Returns `function_group`, `years`, `budgeted`, `spent`, `net`, `worst_year` — for example: `function_group=7400 - Replace Equipment`, `years=5`, `budgeted=728952.0`, `spent=1277394.0`, `net=548442.0`, `worst_year=-0.0103`

What went through fund 1301, and when?

```
``sql SELECT fy, period, eff_date, src_meaning, COUNT(*) AS entries FROM fund_1301_cash_journal GROUP BY fy, period, eff_date, src_meaning ORDER BY fy DESC, eff_date DESC LIMIT 20 ``
```

Returns `fy`, `period`, `eff_date`, `src_meaning`, `entries` — for example: `fy=2026`, `period=12`, `eff_date=2026-06-12`, `src_meaning=payroll journal`, `entries=1`

Which accounts had the largest unspent balance at the latest period?

```
``sql SELECT a.name, a.fund_name, l.fy, l.period, ROUND(l.available) AS available FROM
ledger_snapshot l JOIN account a USING (account_id) ORDER BY l.fy DESC, l.period DESC,
l.available DESC LIMIT 20 ``
```

Returns `name`, `fund_name`, `fy`, `period`, `available` — for example: `name=SPED PRIVA`,
`fund_name=GENERAL FUND`, `fy=2026`, `period=12`, `available=522629.0`

How much was transferred into or out of each account?

```
``sql SELECT a.name, l.fy, l.period, ROUND(l.transfers) AS transfers FROM ledger_snapshot l JOIN
account a USING (account_id) WHERE l.transfers <> 0 ORDER BY ABS(l.transfers) DESC LIMIT 20 ``
```

Returns `name`, `fy`, `period`, `transfers` — for example: `name=TR CAP PRO`, `fy=2026`, `period=12`,
`transfers=1240820.0`

transfers is CUMULATIVE. Movement between two periods is the difference of the column, never the later value.

Which accounts are revenue rather than expenditure?

```
``sql SELECT account_type, level, COUNT(*) AS accounts FROM account GROUP BY account_type, level
ORDER BY accounts DESC ``
```

Returns `account_type`, `level`, `accounts` — for example: `account_type=expense`, `level=account`, `accounts=692`

Revenue rows are stored NEGATIVE, exactly as MUNIS prints them. Check `account_type` before doing arithmetic across types.

What share of its budget had each department used?

```
``sql SELECT a.dept, a.name, l.fy, l.period, ROUND(l.pct_used,1) AS pct_used FROM
ledger_snapshot l JOIN account a USING (account_id) WHERE l.pct_used IS NOT NULL ORDER BY
l.pct_used DESC LIMIT 20 ``
```

Returns `dept`, `name`, `fy`, `period`, `pct_used` — for example: `dept=None`, `name=PS TUITION`, `fy=2026`,
`period=9`, `pct_used=752.8`

Staff on the rosters

How many people of each kind did the town print on a roster, by year?

```
``sql SELECT fy, role_category, COUNT(*) AS people FROM v_staff_roster WHERE role_category <>
'unknown' GROUP BY fy, role_category ORDER BY fy, people DESC ``
```

Returns `fy`, `role_category`, `people` — for example: `fy=2011`, `role_category=teacher`, `people=69`

A count of names the town printed. It carries no FTE and no funding source, so it is not a staffing level.

How many paraprofessionals, by school and year?

```
``sql SELECT fy, school, COUNT(*) AS paras FROM v_staff_roster WHERE
role_category='paraprofessional' GROUP BY fy, school ORDER BY fy, paras DESC ``
```

Returns `fy`, `school`, `paras` — for example: `fy=2011`, `school=primary`, `paras=16`

How many paraprofessionals were tied to a named grade?

```
``sql SELECT fy, role_grade, COUNT(*) AS paras FROM v_staff_roster WHERE
role_category='paraprofessional' AND role_grade <> '' GROUP BY fy, role_grade ORDER BY fy,
role_grade ``
```

Returns `fy`, `role_grade`, `paras` — for example: `fy=2011`, `role_grade=2`, `paras=2`

What did the town print as the title for a paraprofessional, year by year?

```
``sql SELECT fy, role_raw, COUNT(*) AS rows FROM v_staff_roster WHERE
role_category='paraprofessional' AND role_raw <> '' GROUP BY fy, role_raw ORDER BY fy, rows DESC
``
```

Returns `fy`, `role_raw`, `rows` — for example: `fy=2011`, `role_raw=Tutors/Aides`, `rows=17`

Tutor, Aide, Tutors/Aides, Paraprofessional, Para, (para), Sped Para. Five names for one job across fifteen years, which is why `role_category` exists.

Which roster titles could not be classified at all?

```
``sql SELECT role_raw, grade_or_dept, rows FROM role_classification WHERE
role_category='unknown' ORDER BY rows DESC LIMIT 20 ``
```

Returns `role_raw`, `grade_or_dept`, `rows` — for example: `role_raw=`, `grade_or_dept=Extended Day`, `rows=8`

Left `unknown` rather than guessed. 8% of rows.

Which rule decided each classification, and how much rests on the weakest ones?

```
``sql SELECT classified_by, role_category, SUM(rows) AS rows FROM role_classification GROUP BY
classified_by, role_category ORDER BY rows DESC ``
```

Returns `classified_by`, `role_category`, `rows` — for example: `classified_by=heading-department`, `role_category=teacher`, `rows=1103`

A rule beginning `heading-` read the section heading rather than a printed title, which is weaker evidence.

How many names appear on each school roster in each year?

```
``sql SELECT fy, school, position, count FROM staff_roster_counts ORDER BY fy DESC, count DESC
LIMIT 20 ``
```

Returns `fy`, `school`, `position`, `count` — for example: `fy=2025`, `school=turkey-hill`, `position=(unmapped)`, `count=6`

How many teachers were printed against a specific grade?


```
``sql SELECT fy, role_grade, COUNT(*) AS teachers FROM v_staff_roster WHERE
role_category='teacher' AND role_grade <> '' GROUP BY fy, role_grade ORDER BY fy DESC,
role_grade ``
```

Returns `fy`, `role_grade`, `teachers` — for example: `fy=2025`, `role_grade=1`, `teachers=5`

Which schools have roster entries, and for which years?

```
``sql SELECT school, COUNT(DISTINCT fy) AS years, MIN(fy) AS first, MAX(fy) AS last, COUNT(*) AS
rows FROM staff_roster_entries GROUP BY school ORDER BY rows DESC ``
```

Returns `school`, `years`, `first`, `last`, `rows` — for example: `school=high`, `years=15`, `first=2011`, `last=2025`, `rows=1070`

How many counselors, nurses, psychologists and social workers, by year?

```
``sql SELECT fy, role_category, COUNT(*) AS people FROM v_staff_roster WHERE role_category IN
('counselor','nurse','psychologist','social_worker','speech_therapist') GROUP BY fy,
role_category ORDER BY fy, role_category ``
```

Returns `fy`, `role_category`, `people` — for example: `fy=2011`, `role_category=counselor`, `people=9`

How many people did the town print on a roster in total, by year?

```
``sql SELECT fy, COUNT(*) AS names, COUNT(DISTINCT school) AS schools FROM v_staff_roster GROUP
BY fy ORDER BY fy ``
```

Returns `fy`, `names`, `schools` — for example: `fy=2011`, `names=216`, `schools=5`

Which names appear across the most years?

```
``sql SELECT name, COUNT(DISTINCT fy) AS years, MIN(fy) AS first, MAX(fy) AS last FROM
staff_roster_entries WHERE name <> '' GROUP BY name ORDER BY years DESC, name LIMIT 20 ``
```

Returns `name`, `years`, `first`, `last` — for example: `name=Erin Blanchette`, `years=15`, `first=2011`, `last=2025`

| *A name printed in a public annual report. It is not a claim about employment, and a roster gives no FTE.*

How many administrators did each school print?

```
``sql SELECT fy, school, COUNT(*) AS admins FROM v_staff_roster WHERE
role_category='administrator' GROUP BY fy, school ORDER BY fy DESC, admins DESC LIMIT 20 ``
```

Returns `fy`, `school`, `admins` — for example: `fy=2025`, `school=middle`, `admins=5`

Which grades appear anywhere on the rosters?

```
``sql SELECT role_grade, COUNT(*) AS rows FROM v_staff_roster WHERE role_grade <> '' GROUP BY
role_grade ORDER BY rows DESC ``
```

Returns `role_grade`, `rows` — for example: `role_grade=K`, `rows=140`

Athletics and fees

What did each sport cost, and how many played?

```
``sql SELECT fy, season, level, sport, metric, value FROM athletics_by_sport WHERE is_numeric='1' ORDER BY fy DESC, value DESC LIMIT 20 ``
```

Returns `fy`, `season`, `level`, `sport`, `metric`, `value` — for example: `fy=2026`, `season=Spring`, `level=MS`, `sport=Track`, `metric=Total Athletes`, `value=46.0`

Do the per-sport figures add up to the totals the district printed?

```
``sql SELECT season, scope, metric, fy, printed, summed_from_rows, difference, ties FROM athletics_by_sport_reconciliation ORDER BY ABS(CAST(difference AS REAL)) DESC LIMIT 20 ``
```

Returns `season`, `scope`, `metric`, `fy`, `printed`, `summed_from_rows`, `difference`, `ties` — for example: `season=Fall`, `scope=ALL`, `metric=Full Pay`, `fy=2024`, `printed=45050.0`, `summed_from_rows=197.0`, `difference=44853.0`, `ties=0`

What has athletics cost and raised, year by year?

```
``sql SELECT fy, side, item, amount, basis FROM athletics_history ORDER BY fy DESC, side ``
```

Returns `fy`, `side`, `item`, `amount`, `basis` — for example: `fy=2026`, `side=general`, `item=Athletic Coaches`, `amount=159444.0`, `basis=budget`

What has the athletic fee been, by year and tier?

```
``sql SELECT fy, school_year, level, item, amount, unit, verified FROM athletic_fee_schedule ORDER BY fy DESC, level ``
```

Returns `fy`, `school_year`, `level`, `item`, `amount`, `unit`, `verified` — for example: `fy=2027`, `school_year=2026-27`, `level=ANY`, `item=family_cap`, `amount=1500.00`, `unit=per student per sport per season`, `verified=source not held`

Which rates does this project know about, and which does it use?

```
``sql SELECT fy, category, item, value, value_type, set_on FROM rate_register ORDER BY fy DESC, category LIMIT 20 ``
```

Returns `fy`, `category`, `item`, `value`, `value_type`, `set_on` — for example: `fy=2029`, `category=contract_cola`, `item=cost-of-living adjustment to the salary scale`, `value=2.5`, `value_type=percent`, `set_on=`

It deliberately includes rates the model does NOT use, and the ones that cannot be stated at all.

Which rates were set by a document we hold, and which were not?

```
``sql SELECT category, COUNT(*) AS rates, SUM(CASE WHEN source_file <> '' THEN 1 ELSE 0 END) AS with_a_document FROM rate_register GROUP BY category ORDER BY rates DESC ``
```

Returns `category`, `rates`, `with_a_document` — for example: `category=athletic_fee`, `rates=28`, `with_a_document=24`

The annual town reports

What appropriations did each annual report print, and are the rows checked?

```
``sql SELECT edition, status, COUNT(*) AS rows FROM report_appropriations GROUP BY edition, status ORDER BY edition, rows DESC ``
```

Returns `edition`, `status`, `rows` — for example: `edition=FY2011`, `status=check failed`, `rows=298`

ALWAYS split on status. checked, check failed and no check are three different claims and nothing may be aggregated across them.

What did the town pay in gross wages, and to how many people?

```
``sql SELECT edition, status, COUNT(*) AS rows FROM report_gross_wages GROUP BY edition, status ORDER BY edition ``
```

Returns `edition`, `status`, `rows` — for example: `edition=FY2011`, `status=no check`, `rows=326`

What debt has the town carried?

```
``sql SELECT edition, COUNT(*) AS rows, SUM(CASE WHEN status='checked' THEN 1 ELSE 0 END) AS checked FROM report_debt GROUP BY edition ORDER BY edition ``
```

Returns `edition`, `rows`, `checked` — for example: `edition=FY2011`, `rows=20`, `checked=0`

What capital projects did the reports list?

```
``sql SELECT edition, label, status FROM report_capital_projects WHERE label <> '' ORDER BY edition DESC LIMIT 20 ``
```

Returns `edition`, `label`, `status` — for example: `edition=FY2025`, `label=3006 8/2STM & 6/6 STM Dev Cem`, `status=no check`

What is in the trust funds?

```
``sql SELECT edition, COUNT(*) AS rows FROM report_trust_funds GROUP BY edition ORDER BY edition ``
```

Returns `edition`, `rows` — for example: `edition=FY2011`, `rows=66`

What did the town value its property at?

```
``sql SELECT edition, label, status FROM report_valuation WHERE label <> '' ORDER BY edition DESC LIMIT 20 ``
```

Returns `edition`, `label`, `status` — for example: `edition=FY2019`, `label=_Personal Property`, `status=no check`

What enrollment and MCAS results were printed?

```
``sql SELECT edition, COUNT(*) AS rows FROM report_enrollment_mcas GROUP BY edition ORDER BY edition ``
```

Returns `edition`, `rows` — for example: edition=FY2011, rows=5

What did the town assess for Monty Tech?

```
``sql SELECT edition, label, status FROM report_monty_tech WHERE label <> '' ORDER BY edition DESC LIMIT 20 ``
```

Returns `edition`, `label`, `status` — for example: edition=FY2017, label=Chapter 70 13,764,000, status=no check

How many births, deaths and marriages were recorded?

```
``sql SELECT edition, label, status FROM report_vital_records WHERE label <> '' ORDER BY edition DESC LIMIT 20 ``
```

Returns `edition`, `label`, `status` — for example: edition=FY2025, label=Annual Town Meeting - May, status=no check

Who held town office, and when?

```
``sql SELECT edition, label FROM report_officials WHERE label <> '' ORDER BY edition DESC LIMIT 20 ``
```

Returns `edition`, `label` — for example: edition=FY2023, label=Unemployment

What did each department report doing?

```
``sql SELECT edition, COUNT(*) AS rows FROM report_dept_activity GROUP BY edition ORDER BY edition ``
```

Returns `edition`, `rows` — for example: edition=FY2011, rows=3

What receipts did the town record, and from what source?

```
``sql SELECT fy, source, amount, status FROM annual_report_receipts ORDER BY fy DESC, CAST(amount AS REAL) DESC LIMIT 20 ``
```

Returns `fy`, `source`, `amount`, `status` — for example: fy=2023, source=TRAILER PAARE, amount=14124.00, status=no check

What tables does each annual report contain?

```
``sql SELECT fy, [table], pages, figure_rows, checkable FROM annual_report_contents ORDER BY fy DESC, figure_rows DESC LIMIT 20 ``
```

Returns `fy`, `table`, `pages`, `figure_rows`, `checkable` — for example: fy=2025, table=enterprise, pages=25,35,53,64,67,140,142-143,152, figure_rows=96, checkable=no

Which tables in the reports were read, and which are still uncaptured?

```
``sql SELECT dataset, COUNT(*) AS editions, SUM(CASE WHEN extractable='yes' THEN 1 ELSE 0 END) AS extractable FROM extraction_plan GROUP BY dataset ORDER BY editions DESC LIMIT 20 ``
```

Returns dataset, editions, extractable — for example: dataset=enrollment_mcas, editions=94, extractable=0

What did the survey find on each page of each report?

```
``sql SELECT fy, mode, COUNT(*) AS pages, SUM(money) AS money_tokens FROM annual_report_survey GROUP BY fy, mode ORDER BY fy, pages DESC ``
```

Returns fy, mode, pages, money_tokens — for example: fy=2011, mode=ocr, pages=101, money_tokens=2304

What is catalogued in each report, by printed heading?

```
``sql SELECT fy, printed_heading, pages, grain FROM annual_report_catalogue WHERE printed_heading <> '' ORDER BY fy DESC LIMIT 20 ``
```

Returns fy, printed_heading, pages, grain — for example: fy=2025, printed_heading=LUNENBURG PROFILE, pages=9, grain=One statistic

Where did the extraction find something it could not reconcile?

```
``sql SELECT fy, edition, [table], kind, detail FROM report_anomalies ORDER BY fy DESC LIMIT 20 ``
```

Returns fy, edition, table, kind, detail — for example: fy=2025, edition=FY2025, table=FY2026 program of capital projects, Option 1, kind=announced but absent, detail=CPC rankings are not contiguous — 14, 18 and 21 are absent from the list, so it is a filtered ranking, not a complete one.

| *An anomaly is a finding about our reading of the page as much as about the page.*

Which special revenue funds appear in the reports, and in which years?

```
``sql SELECT fy, [group], COUNT(*) AS rows FROM special_revenue_funds GROUP BY fy, [group] ORDER BY fy DESC, rows DESC LIMIT 20 ``
```

Returns fy, group, rows — for example: fy=2025, group=, rows=177

How many figures did each annual report yield, and how many were checked?

```
``sql SELECT edition, COUNT(*) AS rows, SUM(CASE WHEN status='checked' THEN 1 ELSE 0 END) AS checked FROM report_appropriations GROUP BY edition ORDER BY edition ``
```

Returns edition, rows, checked — for example: edition=FY2011, rows=298, checked=0

Which report tables print a total we can reconcile to?

```
``sql SELECT fy, [table], printed_total, checkable FROM annual_report_contents WHERE printed_total <> '' ORDER BY fy DESC LIMIT 20 ``
```

Returns `fy`, `table`, `printed_total`, `checkable` — for example: `fy=2025`, `table=capital_project`, `printed_total=TOTAL CAPITAL PROJECT FUND BALANCE=3,096,913.16`, `checkable=yes`

How many pages of each report carried figures at all?

```
``sql SELECT fy, COUNT(*) AS pages, SUM(CASE WHEN money > 0 THEN 1 ELSE 0 END) AS with_money
FROM annual_report_survey GROUP BY fy ORDER BY fy ``
```

Returns `fy`, `pages`, `with_money` — for example: `fy=2011`, `pages=101`, `with_money=44`

What kinds of anomaly did the extraction find, and how often?

```
``sql SELECT kind, COUNT(*) AS occurrences FROM report_anomalies GROUP BY kind ORDER BY
occurrences DESC ``
```

Returns `kind`, `occurrences` — for example: `kind=unreadable / OCR damage`, `occurrences=167`

Revenue, tax base and free cash

What free cash has each town certified, and from what?

```
``sql SELECT town, year, line, amount, role FROM free_cash_proof ORDER BY year DESC, town LIMIT
20 ``
```

Returns `town`, `year`, `line`, `amount`, `role` — for example: `town=Ayer`, `year=2025`, `line=Free Cash Certified Prior Year`, `amount=3261808.00`, `role=prior_year_certified`

Absolute dollars with no denominator, so they do not compare between towns of different size. The composition does compare, because a share has no size.

How does Lunenburg free cash compare with its neighbours, by composition?

```
``sql SELECT town, year, line, amount FROM free_cash_proof WHERE role='component' ORDER BY year
DESC, town LIMIT 20 ``
```

Returns `town`, `year`, `line`, `amount` — for example: `town=Ayer`, `year=2025`, `line=Revenue Deficits`, `amount=0.00`

What has the town spent on capital, and where did the money come from?

```
``sql SELECT fy, total, free_cash, taxation, unexpended_prior_year_capital, other FROM
capital_funding_history ORDER BY fy ``
```

Returns `fy`, `total`, `free_cash`, `taxation`, `unexpended_prior_year_capital`, `other` — for example: `fy=2017`, `total=619475.00`, `free_cash=250000.00`, `taxation=349023.05`, `unexpended_prior_year_capital=20451.95`, `other=0.00`

What is in the FY27 capital plan, and what is funded?

```
``sql SELECT rank, dept, project, cost, funded, funding FROM capital_plan_fy27 ORDER BY rank ``
```

Returns rank, dept, project, cost, funded, funding — for example: rank=1, dept=DPW, project=Flat Hill Rd Bridge Completion (Engineering and Construction), cost=350000.00, funded=yes, funding=free_cash_or_taxation

How do state measures compare Lunenburg with other districts?

```
``sql SELECT district, fy, measure, value FROM dese_measure WHERE district <> '' ORDER BY fy  
DESC LIMIT 20 ``
```

Returns district, fy, measure, value — for example: district=Harvard, fy=2025, measure=In-District FTE Pupils, value=1016.4

Which DESE measures reconcile against the printed totals, and which do not?

```
``sql SELECT measure, COUNT(*) AS rows, SUM(CASE WHEN reconciles='1' THEN 1 ELSE 0 END) AS  
reconciling FROM dese_measure GROUP BY measure ORDER BY rows DESC LIMIT 15 ``
```

Returns measure, rows, reconciling — for example: measure=Total In-District Expenditures, rows=102, reconciling=0

How has free cash moved for Lunenburg specifically?

```
``sql SELECT year, line, amount, role FROM free_cash_proof WHERE town='Lunenburg' ORDER BY year  
DESC, line ``
```

Returns year, line, amount, role — for example: year=2025, line=Add Actual Revenue Received but not Estimated (CL#7), amount=7139.00, role=component

Which towns does the free cash comparison cover?

```
``sql SELECT town, COUNT(DISTINCT year) AS years, MIN(year) AS first, MAX(year) AS last FROM  
free_cash_proof GROUP BY town ORDER BY town ``
```

Returns town, years, first, last — for example: town=Ayer, years=5, first=2021, last=2025

What measures does the state publish about this district?

```
``sql SELECT measure, COUNT(*) AS rows, MIN(fy) AS first, MAX(fy) AS last FROM dese_measure  
GROUP BY measure ORDER BY rows DESC LIMIT 15 ``
```

Returns measure, rows, first, last — for example: measure=Total In-District Expenditures, rows=102, first=2009, last=2025

Votes and elections

What has the town been asked to fund, and did it agree?

```
``sql SELECT date, election, question, type, amount, yes, no, total FROM ballot_questions ORDER  
BY date DESC ``
```

Returns date, election, question, type, amount, yes, no, total — for example: date=2025-05, election=Annual Town Meeting, question=ARTICLE 11 (Citizens Petition), type=Proposition 2½ override, amount=2099337, yes=, no=, total=

What turnout did each ballot question draw?

```
``sql SELECT date, question, total, registered, ROUND(100.0*total/registered,1) AS turnout_pct
FROM ballot_questions WHERE registered > 0 ORDER BY date DESC ``
```

Returns date, question, total, registered, turnout_pct — for example: date=2014-01-11, question=QUESTION 1. DEBT EXCLUSION, total=1957, registered=7059, turnout_pct=27.7

What election results did the annual reports print?

```
``sql SELECT edition, COUNT(*) AS rows FROM report_elections GROUP BY edition ORDER BY edition
``
```

Returns edition, rows — for example: edition=FY2011, rows=16

Provenance, and what is not established

Where did a figure in this dataset come from?

```
``sql SELECT dataset, edition, document, publisher_label, sha256 FROM dataset_document ORDER BY
dataset, edition LIMIT 20 ``
```

Returns dataset, edition, document, publisher_label, sha256 — for example: dataset=annual-report-catalogue, edition=FY2011, document=4117-fy-2011-annual-town-report.pdf, publisher_label=FY 2011 Annual Town Report, sha256=551f9edd75d051ae4d863a5d5c13925cob9a2f93459e63268c507c5226398b81

This is the join that gives an annual-report row an address. Use it in any query whose answer somebody might want to check.

Which documents does the archive hold, and how were they obtained?

```
``sql SELECT source_type, basis, COUNT(*) AS documents FROM document GROUP BY source_type, basis
ORDER BY documents DESC ``
```

Returns source_type, basis, documents — for example: source_type=primary, basis=None, documents=294

Which documents no longer open at the publisher, or no longer match our copy?

```
``sql SELECT doc_id, link_state, copy_state, url FROM document WHERE copy_state NOT IN
('identical','') OR link_state NOT IN ('200','') LIMIT 20 ``
```

Returns doc_id, link_state, copy_state, url — for example: doc_id=sources/district-budget/docs/fy26-superintendent-39-s-proposed-budget-2-26-25-updated-3-12-25.pdf, link_state=200, copy_state=reflowed, url=https://drive.google.com/file/d/1sqlWrNsH43AE8JqAAAnqNPpOt1mi3gCDU/view?usp=drive_link

Which documents have no upstream address at all?

```
``sql SELECT doc_id, source_type FROM document WHERE url IS NULL OR url='' LIMIT 20 ``
```

Returns `doc_id`, `source_type` — for example: `doc_id=sources/budget-workbooks/fy27-budget-projection-2-24-26.xlsx`, `source_type=restatement`

A gap on our side, not the town's: they were gathered before the mirror existed and nobody wrote down where they came from.

Where do two sources state the same budget line differently?

```
``sql SELECT label, fy, stage, source, value, is_kept FROM line_history_disagreements ORDER BY fy DESC, label LIMIT 20 ``
```

Returns `label`, `fy`, `stage`, `source`, `value`, `is_kept` — for example: `label=Admin Tech Contracted Services`, `fy=2027`, `stage=proposed`, `source=fy27-budget-projections-as-of-2-16-26-with-restorations.txt`, `value=138202`, `is_kept=0`

How much of each dataset has been checked against the page it came from?

```
``sql SELECT dataset, reconciled, partial, SUM(CAST(rows AS INTEGER)) AS rows FROM dataset_document GROUP BY dataset, reconciled, partial ORDER BY rows DESC LIMIT 20 ``
```

Returns `dataset`, `reconciled`, `partial`, `rows` — for example: `dataset=report-appropriations`, `reconciled=0`, `partial=0`, `rows=4665`

Which figures has somebody stated publicly, and on what basis?

```
``sql SELECT fy, metric, amount, stated_by, stated_on, basis FROM stated_figure ORDER BY fy DESC ``
```

Returns `fy`, `metric`, `amount`, `stated_by`, `stated_on`, `basis` — for example: `fy=2025`, `metric=school_surplus`, `amount=603885.97`, `stated_by=Business Administrator (Mr. McNamara) to the School Committee`, `stated_on=2025-09-17`, `basis=close`

Which budget lines have no ledger account mapped to them?

```
``sql SELECT COUNT() AS budget_lines, (SELECT COUNT() FROM crosswalk) AS mapped FROM budget_line ``
```

Returns `budget_lines`, `mapped` — for example: `budget_lines=688`, `mapped=0`

The crosswalk is empty ON PURPOSE. District lines are named, MUNIS rows are coded, and no published document maps one to the other. Budget-to-actual at line level cannot be answered from this data.

Which pages of the annual reports have columns we could not establish?

```
``sql SELECT edition, COUNT(*) AS rows FROM report_appropriations WHERE column_meaning LIKE 'not established%' GROUP BY edition ORDER BY rows DESC ``
```

Returns `edition`, `rows` — for example: `edition=FY2014`, `rows=90`

v1 is an ordinal -- the first column of THIS page that held figures -- not a column name. Read `column_meaning` before summing anything.

How many rows of each report table failed their own check?

```
``sql SELECT 'appropriations' AS t, status, COUNT() AS rows FROM report_appropriations GROUP BY status UNION ALL SELECT 'gross_wages', status, COUNT() FROM report_gross_wages GROUP BY status ORDER BY t, rows DESC ``
```

Returns `t`, `status`, `rows` — for example: `t=appropriations`, `status=check failed`, `rows=4530`

Which documents were obtained by records request rather than published?

```
``sql SELECT source_type, COUNT(*) AS documents FROM document GROUP BY source_type ORDER BY documents DESC ``
```

Returns `source_type`, `documents` — for example: `source_type=primary`, `documents=294`

What basis does each document have for the figures it prints?

```
``sql SELECT basis, COUNT(*) AS documents FROM document WHERE basis IS NOT NULL GROUP BY basis ORDER BY documents DESC ``
```

Returns `basis`, `documents` — for example: `basis=budget`, `documents=173`

ledger means a figure exists because a transaction did. restatement means a prior year re-presented by the party that spent it. They are not interchangeable.

Which datasets have a document for every edition, and which do not?

```
``sql SELECT dataset, COUNT(*) AS editions, SUM(CASE WHEN document <> '' THEN 1 ELSE 0 END) AS with_a_document FROM dataset_document GROUP BY dataset ORDER BY editions DESC ``
```

Returns `dataset`, `editions`, `with_a_document` — for example: `dataset=annual-report-catalogue`, `editions=16`, `with_a_document=16`

How many rows of each dataset came off each page?

```
``sql SELECT dataset, edition, pages, rows FROM dataset_document ORDER BY CAST(rows AS INTEGER) DESC LIMIT 20 ``
```

Returns `dataset`, `edition`, `pages`, `rows` — for example: `dataset=report-gross-wages`, `edition=FY2025`, `pages=177,178,179,180,181,182`, `rows=607`

Finding your way

What tables are there, and how big are they?

```
``sql SELECT name FROM sqlite_master WHERE type IN ('table','view') ORDER BY type, name ``
```

Returns `name` — for example: `name=account`

What fiscal years does the archive cover, per dataset?

```
``sql SELECT dataset, MIN(edition) AS first, MAX(edition) AS last, COUNT(*) AS editions FROM dataset_document GROUP BY dataset ORDER BY dataset ``
```

Returns `dataset`, `first`, `last`, `editions` — for example: `dataset=annual-report-catalogue`, `first=FY2011`, `last=FY2025`, `editions=16`

What does one fiscal period mean?

```
``sql SELECT period, label, months_elapsed, is_final FROM fiscal_period ORDER BY period ``
```

Returns `period`, `label`, `months_elapsed`, `is_final` — for example: `period=1`, `label=July`, `months_elapsed=1.0`, `is_final=0`