

Is my ad wearing out? Run it on your own data

A brief to hand to your AI assistant - Stochastic Strata

For you (the human) - read this first

This is a little technical: it means exporting a table from your ad platform and running a free notebook. You do NOT have to do it alone. Give this whole document to a capable AI assistant - ChatGPT, Claude, Gemini or similar - by uploading the PDF or pasting its text, then say: "Help me do this, step by step." Everything the assistant needs is below, even if you have never written a line of code.

What you get at the end is a verdict for each creative, and the method names how much it could see. Wearing out, fitted: the ad's response is falling as people see it more, and the notebook reads off where it was performing before it tired, at what exposure it started to tire, and how fast it is losing what is left - each as a range. Wearing out, range only: the drop is real and exposure rose, but the data could not support the full reading, so there is no onset and no rate. Decline, cause unknown: the drop is real but exposure did not move, so something other than the creative changed. Steady: nothing separates from ordinary wobble. Cannot tell: too little volume or history to judge, which is an honest refusal rather than a clean bill of health.

One thing only you can supply: a reach or frequency column. Almost every export gives impressions and clicks; far fewer give the number of PEOPLE reached to date. That column is what builds the axis the whole method runs on - how many times the average person has seen the ad. Without it you can still see whether performance fell, but nothing can tell you whether the cause was fatigue - and that is the whole question.

The post, the notebook, the spreadsheet and the interactive version are here:

<https://www.stochasticstrata.com/blog/is-my-ad-wearing-out>

Instructions for the AI assistant

You are guiding a marketer through deciding whether an ad creative is wearing out. The method puts the creative's daily click-through rate on the EXPOSURE axis - impressions delivered so far divided by distinct people reached so far - and fits a settled level, an onset and a decay to that curve, reporting a bend only when it clearly beats a flat line. When it cannot fit a bend it steps down: it compares the last week against a settled early window, asks where the week would land if nothing had changed, and requires exposure to have risen MATERIALLY before it will call the decline fatigue. Guide them ONE STEP AT A TIME, in plain language, at their level. Do not assume they know SQL, BigQuery or Python.

Start by asking what they can export, and specifically whether they can get reach or frequency alongside impressions and clicks. If they cannot, say plainly what that costs them: there is no axis, the bend cannot be fitted, and the analysis degrades to a decline detector that cannot attribute a cause. Do not paper over it.

You need the notebook, and here is the measured reason why.

This brief tells you how to GET the data and how to READ the answer, and it also describes the shape of the method. You may be tempted to just implement it from that description. Here is what happens when someone does. A careful rebuild from this text alone, filling in the constants it does not give you with sensible choices, agrees with the real method on 3 of the 4 worked examples published in the post - and disagrees with it on 13 of 210 simulated creatives.

Both halves of that matter, and the first half is the dangerous one. Checking your rebuild against the post's examples will mostly tell you it is correct. What it gets wrong are the borderline creatives, and those are the only ones where the verdict changes what anybody does. The one worked example it missed is the thin one: a rebuild without the method's support rules fits a bend through noise the data cannot carry, and reports an onset and a decay where the real method says it can only see a decline. In the wider comparison the rebuild both invented bends on creatives that never tired and missed declines the real method caught.

One trap in particular, because it is the one an improviser falls into hardest: the exposure test at the fallback

rung has an obvious wrong form. Asking whether the trend is above zero passes on about half of the runs of a creative that fell for an outside reason, since on flat exposure day-to-day noise alone puts the trend above zero half the time; the material-rise floor the method uses passes 2.3% of those same runs instead.

So: download and run the notebook from the page above, and do NOT present your own version's numbers as the method's. If you have already written one and it matches the post's worked examples, that is precisely the false confidence described above, not a validation.

Step 1 - get one row per creative per day

The notebook needs a CSV or .xlsx with five columns, one row per creative per day. Platform downloads work as they come - title lines, Total rows, numbers like 8,563, and Meta's, TikTok's and Google Ads' own column names:

```
creative_id,date,impressions,clicks,reach
ad_44192,2026-08-14,2841,47,18240
ad_44192,2026-08-15,3102,51,20905
ad_44193,2026-08-14,1180,14,9702
ad_44193,2026-08-15,1244,12,10630
```

creative_id is any stable identifier for the ad. date is the day. impressions and clicks are that day's totals. reach is the number of distinct people reached TO DATE, cumulative - a daily count of unique people summed across days is NOT reach, and the notebook's exposure axis will be wrong if it is fed that. A by-day report from Meta or TikTok gives each day's OWN reach and each day's own frequency, not reach to date; the notebook refuses such a column (it falls from day to day, or never exceeds one day's impressions) rather than give a silent wrong verdict. If the export carries frequency TO DATE, have them name that column in FREQUENCY_COLUMN in the notebook's YOUR FILE HERE cell, next to MY_FILE.

Aim for at least three weeks of history per creative and ideally sixty days. A bend needs a few days before it and several after it to be read at all, and the fallback rung needs four weeks, so a creative launched last Tuesday cannot be judged yet, and the notebook will say so rather than guessing.

Step 2 - getting reach to date

No by-day report carries reach to date, and summing or taking the largest daily reach does not make it. It comes from asking for reach over a range that starts at the creative's launch and ends on each day. Meta's Insights API takes a whole list of such ranges in one call; the shape is below - fill in the placeholders for them and explain each one. In Ads Manager the same thing is one export per end day, each from the launch date. If the platform cannot give reach to date at all, the notebook still runs without the reach column: it reports a decline and its size, and says plainly that it cannot separate fatigue from another cause.

```
GET https://graph.facebook.com/<API_VERSION>/act_<AD_ACCOUNT_ID>/insights
?level=ad
&fields=ad_id,impressions,inline_link_clicks,reach
&time_ranges=[{"since":"<LAUNCH>","until":"<LAUNCH>"},
               {"since":"<LAUNCH>","until":"<LAUNCH + 1 day>"},
               ... one range per day, EVERY one starting at <LAUNCH> ...]
```

Each row: date_stop is the day, impressions and clicks are running totals from launch, reach is reach to date. The notebook's from_ranges_to_date() turns the running totals back into each day's own.

If they export from the ad platform's own reporting UI instead, the report they want is a per-ad, per-day breakdown with impressions and clicks; its reach column is each day's own, so leave it out or replace it with reach to date as above. Totals rows and thousands separators are read as they come.

Step 3 - run the notebook

Two routes. Locally, with Python installed, is the private one: nothing leaves their machine. Google Colab needs no installation but is not private - opening the notebook there and uploading their file puts that file on Google's servers. Say this plainly before they choose, and let them decide; do not pick for them.

Either way: open the notebook, run the cells in order, and in the your-data section set MY_FILE in the cell

marked YOUR FILE HERE to their export's filename. The first half of the notebook runs on synthetic data with a known answer, which is there so they can see the method work before trusting it on their own numbers. Do not skip it.

Step 4 - read the answer honestly

The notebook returns one verdict per creative and names the rung it came from. Translate them like this, and resist the pull toward a cleaner story than the data supports.

- WEARING OUT, FITTED - the response falls along the exposure axis from a fitted onset. Read out the level, the onset and the decay WITH their ranges, never as bare numbers. New creative is the reasonable next move. This is still a screen, not proof: see the caveats.
- WEARING OUT, RANGE ONLY - the last week fell below where it would land unchanged while exposure rose materially, but no bend could be supported. Report the decline and its size; there is no onset and no rate, and do not invent them.
- DECLINE, CAUSE UNKNOWN - performance genuinely fell, but exposure did not move, so there was no axis to fit a bend on. Look at bid changes, budget changes, audience changes and competitor activity BEFORE rebuilding the ad. Replacing the creative here costs weeks and fixes nothing.
- STEADY - no bend beats a flat line, and the movement is within what this ad's own volume produces by chance.
- CANNOT TELL - not enough volume or history to judge. Report this as a refusal. Do not round it down to "no fatigue detected", which is a different and false statement.

The caveats - tell the reader these BEFORE they act

State these up front, not as a footnote after the verdict.

- Exposure is the coordinate, not the cause. The method says the response fell AS exposure accumulated; it does not prove repetition did it. A creative whose response never falls fits no bend, however high its frequency runs.
- A rising-exposure decline can still have another cause. Anything that happens to move together with exposure - the reachable audience getting worse as it exhausts, a bid strategy pulling in colder traffic - looks exactly like fatigue. In the post's own simulation 2.3% of outside-cause declines pass the fallback gate even with it working, and that share is noise in the exposure figures, not such a cause.
- No bend found is not the same as healthy. A creative that never accumulated much exposure has no bend to find, and the method correctly says so. Thirty-one real campaigns in a public dataset produced exactly that - none of them ever passed 3.3 impressions per person. Read it as "wear-out was not testable yet", not as "this ad is fine". Run through the notebook's second bend, 15 of the same 31 showed a dip and a recovery that one bend cannot express, so the shape matters as well as the exposure.
- The refresh-versus-retire split does not come from your data. The published evidence that extra exposures can actively hurt is one study, which its authors call unplanned, against earlier work that found none - and a study where exposures were served effectively at RANDOM, because ad allocation targeted publishers rather than people. A live account is the opposite by construction. Treat any retire threshold as a business call about how much decline to tolerate, not a statistical finding.
- Ads recover, and ads level off, and the model has one bend. A creative that dipped and came back reads as tired if it is scored during the dip. A creative that dropped fast and then HELD at a lower level is the shape the method measures as its own limit: the fit finds a bend on only a small share of those. If a verdict is surprising, look at the whole curve before acting on it.
- Auto-rotated placements cannot be scored this way at all. Where the platform mixes assets inside a group
 - Performance Max, Advantage+ - per-creative exposure is not cleanly separable, so there is no axis. Exclude them and say why.
- Single-platform exports only describe that platform. Someone who sees an ad on two networks is two

different people as far as these numbers are concerned, so exposure is understated wherever audiences overlap.

- This is a screen, not a measurement. It reads a shape and infers a cause. The only thing that settles it is holding the creative constant in a randomised slice of the audience and comparing. What the screen is good for is telling them which two creatives out of forty are worth that test.

If the notebook refuses

The notebook is built to refuse rather than guess, and it will do so in several places: columns it cannot find, more clicks than impressions (usually two exports joined at different date grains), a history too short to hold a bend or the fallback windows, exposure that never moved, or a window under the impression floor. Read those out to the reader as findings. They are the method working, not the method failing.