SEO Technical & On-Page Audit

http://www.tedbaker.com

Hedging Beta has produced this analysis on 22/09/2015.

Index

A) Background and Summary

This document provides a detailed analysis of SEO technical and on-page parameters for the following domain: ***http://www.tedbaker.com.*** It identifies all the parameters that could affect search engine rankings and customer's path journey, and are based on: a) accessibility; b) indexability; and c) on-page rankings factors. Our analysis has focused on a limited number of landing pages and sub-categories, which have been identified through a "top-level navigation" approach that provides a good proxy for the performance of Ted Baker. Please find below a technical and on-page analysis summary, details of which are explained in each section of this document.

Score: Areas where intervention is needed in a scale from 1 (most problematic - immediate attention) to 5 (least problematic) from an SEO perspective. (√=correct √= partially correct/ad-hoc analysis required)

Status: indicates the current performance of the website for the specified parameters.

Description: indicates the analysed SEO parameters, which are identified through the crawling.

Images source: Screaming Frog Tool

SCORE	STATUS	ITEM	DESCRIPTION
1	√	**Robots.txt and Robots Meta Tags**	The robots.txt file has been found on the default location.
1	√	**XML Sitemap**	The sitemap.xml file has not been found on the website root folder, however its location could be different if such specified in Google Search Console.
1	√	**HTTP Status Codes**	4xx have been found and should be corrected whereas necessary. 3xx should be analysed based on mapping objectives.
2	√	**URLs**	Non-optimised URLs have been found (underscore, uppercase, parameters, too long, etc.).
2	√	**Canonical**	The majority of pages have been canonicalised.
3	√	**Title**	Non-optimised titles have been found (duplicate, too long, too short, etc.).
4	√	**Hierarchy Structure**	Hierarchy (H1, H2, etc.) is not fully optimised (missing, duplicate, etc.).
4	√	**Meta Description**	The majority of meta descriptions are missing.
4	√	**Structured Data Mark-Up**	Additional structured data mark-up could be implemented
5	√	**Meta Keywords**	Almost all keywords are missing.
5	√	**Images**	Non-optimised images have been found (missing alt text, image size over 100kb, etc).

B) Technical and On-Page Analysis

Accessibility & Indexation

Robots.txt and Robots Metatag

The **robots.txt** file identifies the directories and paths that are allowed or disallowed for crawling purposes by specific agents.

- ✓ The robots.txt file has been found on the default location (http://www.tedbaker.com/robots.txt).

```
# For all robots
User-agent: *

# Block access to specific groups of pages

Disallow: /de/cart
Disallow: /de/checkout
Disallow: /de/my-account
Disallow: /de/search?*

Disallow: /fr/cart
Disallow: /fr/checkout
Disallow: /fr/my-account
Disallow: /fr/search?*

Disallow: /ie/cart
Disallow: /ie/checkout
Disallow: /ie/my-account
Disallow: /ie/search?*

Disallow: /neu/cart
Disallow: /neu/checkout
Disallow: /neu/my-account
Disallow: /neu/search?*

Disallow: /nl/cart
Disallow: /nl/checkout
Disallow: /nl/my-account
Disallow: /nl/search?*

Disallow: /oeu/cart
Disallow: /oeu/checkout
Disallow: /oeu/my-account
Disallow: /oeu/search?*

Disallow: /row/cart
Disallow: /row/checkout
Disallow: /row/my-account
Disallow: /row/search?*

Disallow: /seu/cart
Disallow: /seu/checkout
Disallow: /seu/my-account
Disallow: /seu/search?*

Disallow: /uk/cart
Disallow: /uk/checkout
Disallow: /uk/my-account
Disallow: /uk/search?*

# Block access to disallowed sites


Disallow: /oeu
Disallow: /seu
Disallow: /neu

Request-rate: 1/10              # maximum rate is one page every 10 seconds
Crawl-delay: 10                 # 10 seconds between page requests
Visit-time: 0400-0845           # only visit between 04:00 and 08:45 UTC

# Allow search crawlers to discover the sitemap
Sitemap: /sitemap.xml

# Block CazoodleBot as it does not present correct accept content headers
User-agent: CazoodleBot
Disallow: /

# Block MJ12bot as it is just noise
User-agent: MJ12bot
Disallow: /

# Block dotbot as it cannot parse base urls properly
User-agent: dotbot/1.0
Disallow: /

# Block Gigabot
User-agent: Gigabot
Disallow: /
```

Fig.1 Robots Analysis

XML Sitemap

Sitemaps are files providing an easy way to inform search engines with regard to pages available for crawling. The use of the Sitemap protocol does not guarantee that web pages are included in search engines, but helps web crawlers do a better job when crawling your site.

- ✓ The sitemap.xml has not been found on the root folder (http://www.tedbaker.com/sitemap.xml), however its location could be different if such specified in Google Search Console.

Fig.2 Sitemap Analysis

HTTP Status Codes

When a search request to retrieve pages is made to the server, the server itself returns an **HTTP status code** in response to the request. For instance, this happens when a user accesses your page on a browser or when Google crawls the page. This status code provides necessary information about the status of the request. Different status codes (i.e., 4xx and 5xx) and redirections techniques have been analysed.

- ✓ Blocked By Robots.txt 2 (0.01%)
- ✓ Success (2xx) 21846 (74.62%)
- ✓ Redirection (3xx) 1832 (6.26%)
- ☒ Client Error (4xx) 5595 (19.11%)

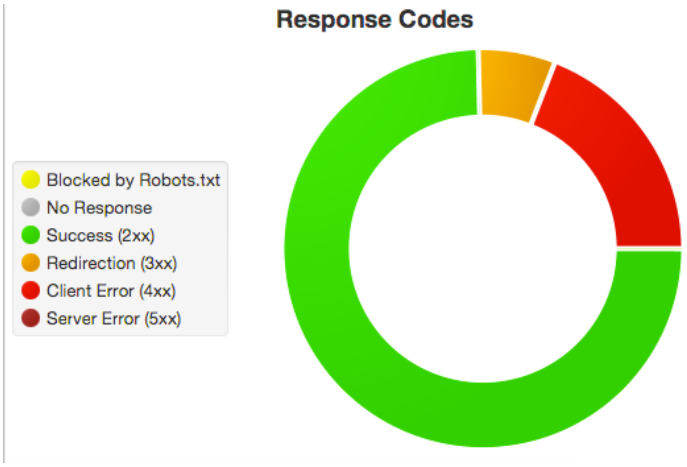

Fig.3 Response Codes Analysis

Page Speed Load

Google has clearly incorporated **site speed** in search rankings, which is why it is becoming increasingly important to analyse the website load speed from a technical point of view. The homepage has been analysed.

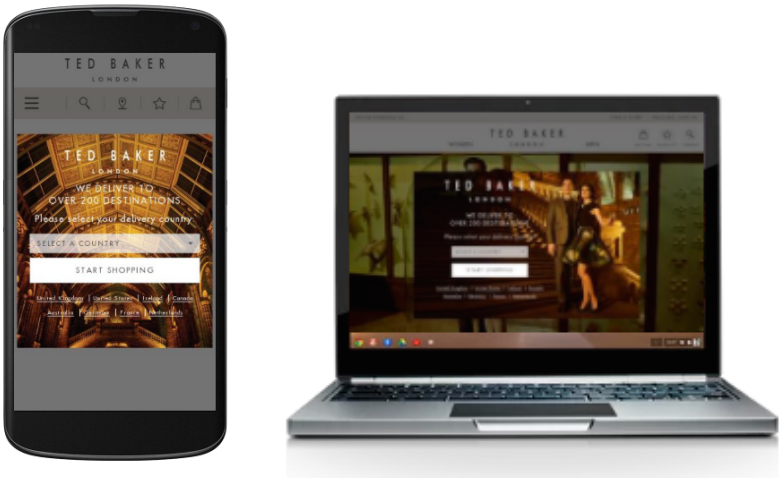

Fig.4 Page Speed Load Analysis

- ✓ Mobile - page speed: 64/100
- ✓ Mobile - user experience 96/100
- ✓ Desktop - overall score: 82/100

Site Architecture

Another important feature is the identification of how many clicks are needed to move away from the homepage to reach other relevant pages and to evaluate how pages are linked in the site's hierarchy.

- ✓ As per graph below, most of the pages are reachable in 5 clicks

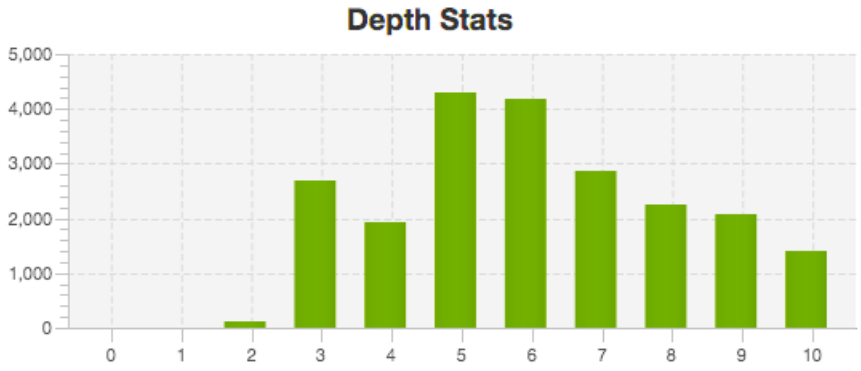

Fig.5 Depth Stats Analysis

On-Page Ranking Factors

The following analysis identifies all the chief characteristics of: a) the site's individual pages; b) the domain.

URLs

Given that a **URL** is the entry point to the content of any page of your website, it's where the on-page analysis begins.
It's important to use URLs that effectively describe their corresponding content, following accurate and appropriate SEO guidelines.

- ☒ 16341 (55.82%) URLs use underscores
- ☒ 27139 (92.70%) URLs use uppercase
- ☒ 16057 (54.85%) URLs use parameters
- ☒ 16883 (57.67%) URLs are over the recommended maximum characters' length

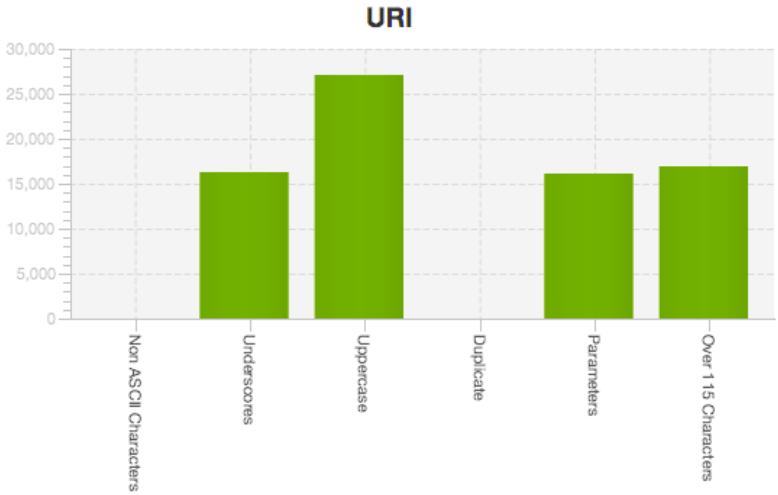

Fig.6 URI Analysis

Title

Unique **titles**, based on targeted keyword and SEO-friendly phrasing, help rankings and performance.

- ☒ 16847 (77.27%) titles are duplicate
- ☒ 6848 (31.41%) titles are below the recommended minimum characters' length
- ☒ 2178 (9.99%) titles are over the recommended maximum characters' length
- ☒ 4179 (19.17%) titles are over the recommended pixels' limit
- ☒ 6848 (31.41%) titles are below the recommended pixels' limit
- ☒ 38 (0.17%) titles are the same as H1s

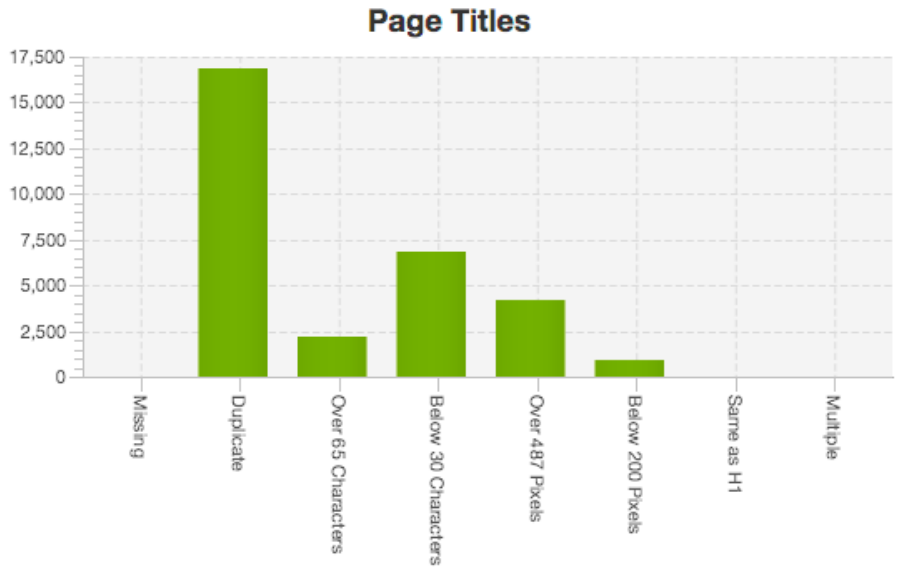

Fig.7 Titles Analysis

Meta Description

It doesn't explicitly act as a ranking factor, but it does affect the page's click-through rate in the search engine results. Having a good **meta description** won't stop Google from choosing a different meta description from the one that has been included (e.g. content or directories' description). However, Google will still consider it as part of "its content analysis", of course.

- ☒ 17731 (81.32%) meta descriptions are missing
- ☒ 4071 (18.67%) meta descriptions are duplicate
- ☒ 2378 (10.91%) meta descriptions are over the recommended maximum characters' length
- ☒ 276 (1.27%) meta descriptions are below the recommended minimum characters' length
- ☒ 2350 (10.78%) meta descriptions are above the recommended pixels' limit
- ☒ 276 (1.27%) meta descriptions are below the recommended pixels' limit

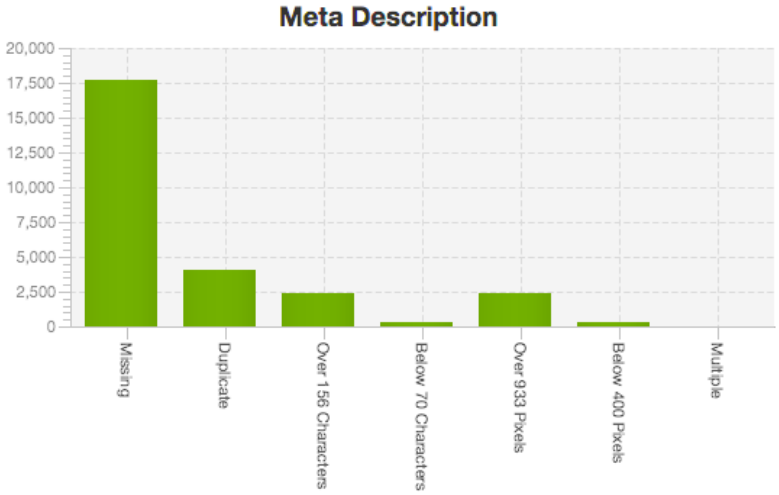

Fig.8 Meta Description Analysis

Meta Keywords

They have become less relevant now than in the last few years, but they still offer residual value, and that's why they must be included in a comprehensive analysis.

- ☒ 21802 (99.99%) meta keywords are missing

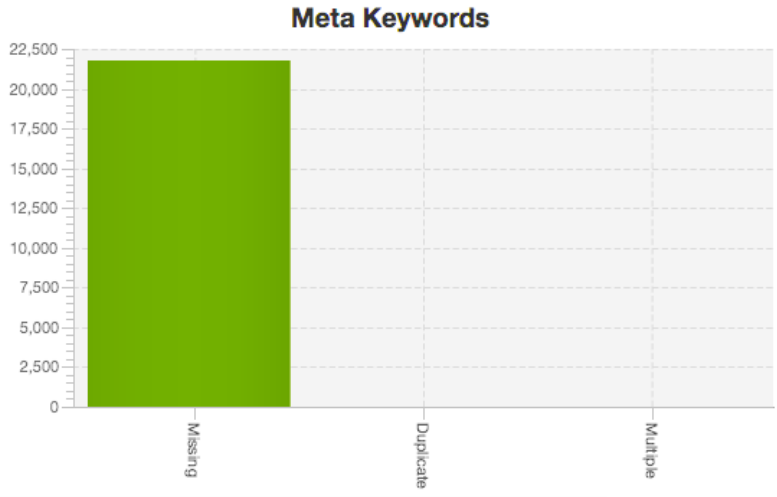

Fig.9 Meta Keywords Analysis

Hierarchy Structure

A clear and defined **hierarchy structure** is important because search engines will use it to determine and possibly carry out the crawling order. Furthermore, the hierarchy structure will help you and the search engine properly understand each element and subject of the website.

- ☒ 406 (1.86%) H1s are missing
- ☒ 20372 (93.43%) H1s are duplicate
- ☒ 1 H2s is missing
- ☒ 21803 (100%) H2s are duplicate
- ☒ 21803 (100%) H2s are multiple

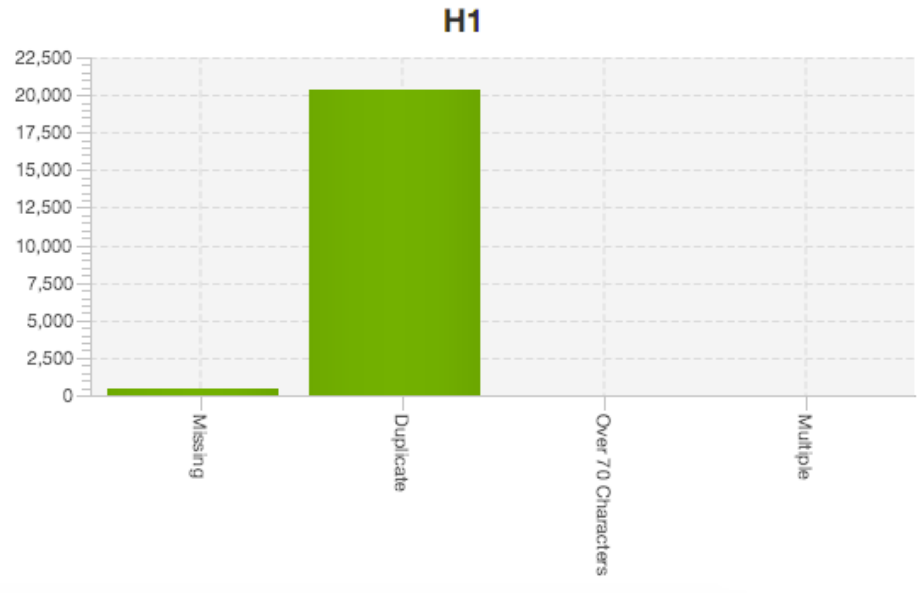

Fig.10 H1s Analysis

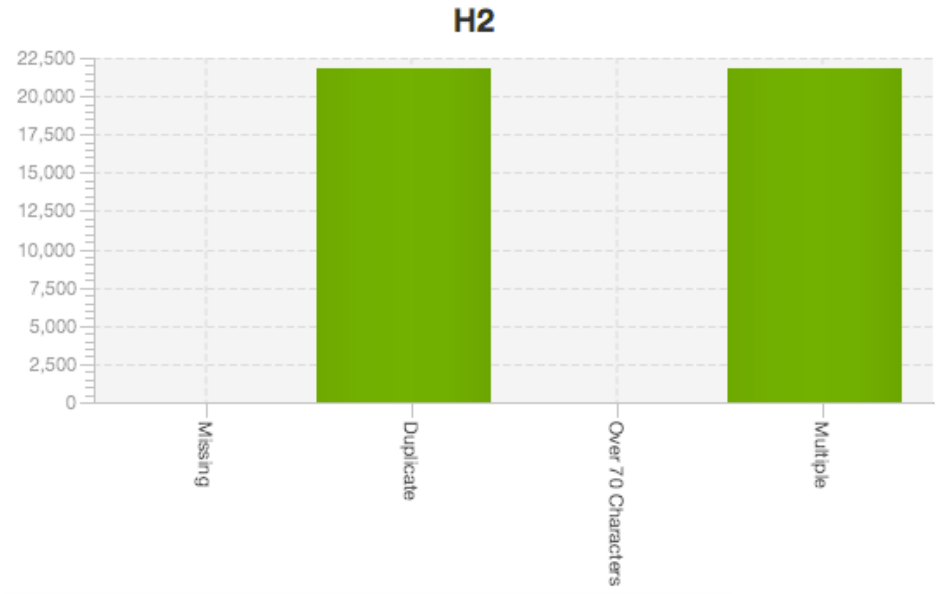

Fig.11 H2s Analysis

Images

It is important to provide **alternative text** (also known as “alt text”) to any images. Otherwise, search engines may not understand the underlying significance of the image itself.

- ☒ 25 (60.98%) Images are over the recommended kb size
- ☒ 16 (39.02%) Images miss an alt text

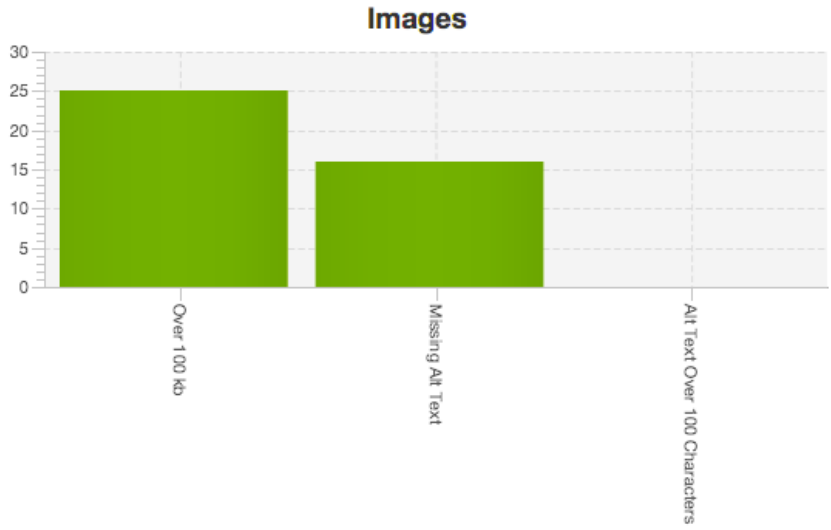

Fig.12 Images Analysis

Rel="canonical" link

The **canonical tag** is used to help avoid duplicate content issues which could lead to website penalisation.

- ✓ 21373 (98.02%) Canonical links have been found
- ✓ 18112 (83.06%) Canonicalised have been found
- ✓ 432 (1.98%) No canonical have been found

Structured data mark-up / Rich Snippets

Rich snippets are part of enhanced Search Engine Results Pages (SERPs) and are designed to help users make decisions and take actions before they click on a specific search result. Moreover, rich snippets help search engines understand the content of your site and differentiate your links from the others. Users are "supported" by rich snippets to go through the discovery of all the sites that are relevant to their searches, which in turn determines higher click-through rates. In addition, rich snippets give webmasters an additional layer of optimisation, while the content of the site is highlighted in the SERPs.

- ✓ Additional structured data mark-up could be implemented