```python
import requests
from bs4 import BeautifulSoup
from urllib.parse import urlparse, urljoin
import random
```

```python
# Function: Get website name from a URL
def get_website_name(url):
    parsed_url = urlparse(url)
    # Get the netloc attribute, containing the domain name
    website_name = parsed_url.netloc
    # Remove 'www.' prefix if present
    if website_name.startswith('www.'):
        website_name = website_name[4:]
    return website_name

# TESTING
URL = "https://www.eurogamer.net/stardew-valley-creator-reveals-haunted-chocolatier"
print(get_website_name(URL))
```

```
eurogamer.net
```

```python
# Function: Counting the number of name in link_set
def count_links_containing_website_name(link_set, name):
    return sum(name in string for string in link_set)

# TESTING
a = {'https://www.eurogamer.net/stardew-valley-creator-says-16-update-content-is-done-definitely-out-this-year'
count_links_containing_website_name(a, "eurogamer.net")
```

```
Out[ ]: 17
```

```python
# Adapted code from https://github.com/kjmazidi/NLP/blob/master/Xtra_Python_Material/Web_Scraping/3%20-%20Web%2
# Instead of writing to a file, append to a list

# Function: Scrape all URLs from an initial URL
def get_urls(url, exclude_links = ['google', 'wiki', 'mod', 'png']):
    # Initialize an empty list to store extracted URLs
    url_list = []

    try:
        # Make a GET request to the provided URL
        r = requests.get(url)

        # Check if the request was successful (status code 200)
        if r.status_code == 200:
            # Parse the HTML of the link
            soup = BeautifulSoup(r.text, 'html.parser')

            # Find all <a> tags containing links
            for link in soup.find_all('a'):
                # Get the value of the 'href' attribute of the <a> tag
                link_str = str(link.get('href'))

                # Check if the link contains certain keywords
                if ('Stardew' in link_str or 'stardew' in link_str) and ('Valley' in link_str or 'valley' in link_str):
                    # Remove unnecessary parameters from the URL (if present)
                    if link_str.startswith('/url?q='):
                        link_str = link_str[7:]

                    # Remove additional parameters after '&' (if present)
                    if '&' in link_str:
                        i = link_str.find('&')
                        link_str = link_str[:i]

                    # Exclude URLs containing 'google' or 'wiki' or 'mod'
                    if link_str.startswith('http') and not any(exclude_link in link_str for exclude_link in exclude_links
                        url_list.append(link_str)
    except:
        print("Could not extract links from", url)

    # Return the list of extracted URLs
    return url_list
```

```python
#Function: Return n random elements from a list
def get_n_random_elements(list, NUMBER_OF_ELEMENTS = 2):
    # If list less than NUMBER_OF_ELEMENTS elements, return the list back
    if len(list) < NUMBER_OF_ELEMENTS:
        return list

    # Select NUMBER_OF_ELEMENTS random elements from the list
    else:
        random_elements = random.sample(list, NUMBER_OF_ELEMENTS)
        return random_elements

# TESTING
test = [1, 2, 3, 4, 5]
```

```
test_elements = get_n_random_elements(test, 2)
print(test_elements)
```

[4, 5]

```python
def crawler(initial_urls, MAX_SIZE = 25, URL_PER_LINK = 3):
    # MAX_SIZE is the number of relevent sites
    # URL_PER_LINK is the number of URLs crawled for each unique site visited
    url_queue = []
    url_queue.extend(starter_urls)

    # This visited links set makes sure that there are no repeats in the sites
    visited_links = set()
    visited_links.update(initial_urls)

    # This visited website set keeps the website name
    # visited_websites_name = set()
    # visited_websites_name.add(get_website_name(initial_url))

    # Manually removed links which resulted in poor data (from experience)
    exclude_links = ['google', 'wiki', 'linkedin', 'facebook', 'mod', 'png', 'jpg', 'jpeg', 'instagram']
    while True:
        if len(visited_links) >= MAX_SIZE:
            return visited_links

        # Case where the queue is empty before exiting the function
        if not url_queue:
            # Add the initial URL back
            url_queue.extend(starter_urls)

        # Pop the top URL of the queue
        current_url = url_queue.pop(0)

        visited_links.add(current_url)
        scraped_urls = get_urls(current_url, exclude_links)

        # Filter out already visited URLs
        # scraped_urls = [url for url in scraped_urls if url not in visited_links]

        n_urls = get_n_random_elements(scraped_urls, URL_PER_LINK)

        for url in n_urls:
            try:
                r = requests.get(url)
                if r.status_code == 200:
                    website_name = get_website_name(url)
                    exclude_links.append(website_name)
                    # Append remaining URLs to the queue
                    url_queue.append(url)
                    print("Added to the list:", url)
            except:
                print("Could not add to the list", url)
```

```python
starter_urls = ["https://en.wikipedia.org/wiki/Stardew_Valley", "https://www.eurogamer.net/stardew-valley-creat
#get_urls(starter_url)

urls = crawler(starter_urls)
```

Added to the list: https://web.archive.org/web/20170705220431/http://www.siliconera.com/2017/02/28/chucklefish-bringing-wargroove-stardew-valley-pocket-rumble-switch/
Added to the list: https://web.archive.org/web/20170718172936/http://www.gamasutra.com/view/news/267563/The_4_years_of_selfimposed_crunch_that_went_into_Stardew_Valley.php
Added to the list: https://www.eurogamer.net/stardew-valley-creator-reveals-haunted-chocolatier?view=comments
Added to the list: https://www.eurogamer.net/stardew-valley-creator-reveals-haunted-chocolatier?view=comments
Added to the list: https://www.eurogamer.net/stardew-valleys-long-awaited-16-update-gets-march-release-date-on-pc
Added to the list: http://www.siliconera.com/2017/01/19/stardew-valley-vita-port-investigation/
Added to the list: http://www.gameinformer.com/b/features/archive/2016/04/28/understanding-autism-through-stardew-valley.aspx
Could not add to the list https://gamerant.com/stardew-valley-greenhouse-guide-unlock-layout-crops/
Added to the list: https://blog.us.playstation.com/2016/12/08/country-life-rpg-stardew-valley-coming-to-ps4-on-december-13/#comment-1193243
Added to the list: https://blog.playstation.com/2016/12/08/country-life-rpg-stardew-valley-coming-to-ps4-on-december-13/comment-page-2/#comments
Added to the list: https://blog.playstation.com/2016/12/08/country-life-rpg-stardew-valley-coming-to-ps4-on-december-13/comment-page-2/#comments
Added to the list: https://www.polygon.com/2018/5/14/17352396/stardew-valley-ps-vita-release-date
Added to the list: https://www.destructoid.com/stardew-valley-is-headed-to-ios-this-month-and-then-android-526438.phtml
Added to the list: https://www.vg247.com/2017/08/01/stardew-valley-multiplayer-details-emerge-beta-test-coming-to-steam-at-the-end-of-2017/
Added to the list: https://www.forbes.com/sites/mattperez/2017/01/03/30-under-30-in-games-2017-the-creators-and-founders-behind-tsm-stardew-valley-and-cards-against-humanity/
Added to the list: https://www.gamesradar.com/stardew-valley-update-16-drops-next-month-concernedape-confirms-now-back-to-the-grind/
Added to the list: http://www.pcgamesn.com/stardew-valley/stardew-valley-a-pc-harvest-moon-like-indie-game-rockets-to-the-top-of-steam-top-sellers-chart
Could not add to the list http://stardewvalley.net/faq/
Added to the list: https://www.king5.com/article/entertainment/television/programs/evening/eric-barone-stardew-valley/281-7a596cc6-3c48-458b-bbda-672919d1c717
Added to the list: https://www.bostonglobe.com/arts/2016/03/17/stardew-valley-you-reap-what-you-sow/OtevYl1D31M5x2KWGClHPK/story.html
Added to the list: https://www.gamespot.com/articles/nintendo-switch-version-of-stardew-valley-releases/1100-6453715/
Added to the list: https://www.pcgamer.com/stardew-valley-multiplayer-out-now-on-pc/
Added to the list: http://www.gamasutra.com/view/news/331910/Stardew_Valley_dev_announces_decision_to_selfpublish.php
Added to the list: https://www.nintendolife.com/news/2023/10/stardew-valley-festival-of-seasons-concert-tour-announced
Added to the list: https://www.pushsquare.com/games/ps4/stardew_valley
Added to the list: https://www.stardewvalleyconcert.com/
Added to the list: https://www.purexbox.com/games/xbox-one/stardew_valley
Added to the list: https://www.familygamingdatabase.com/accessibility/Stardew+Valley
Added to the list: https://www.familygamingdatabase.com/accessibility/Stardew+Valley
Added to the list: https://www.eventpop.me/e/17067/stardew-valley
Added to the list: https://www.etix.com/ticket/p/82673938/stardew-valleyfestival-of-seasons-tampa-tampa-theatre
Added to the list: https://www.melbournerecital.com.au/events/2024/stardew-valley-festival-of-seasons/

```
In [ ]: print("Number of links:", len(urls))
        print("Links:")
        for url in urls:
          print(url)
```

```
Number of links: 25
Links:
https://www.eurogamer.net/stardew-valley-creator-reveals-haunted-chocolatier
https://blog.us.playstation.com/2016/12/08/country-life-rpg-stardew-valley-coming-to-ps4-on-december-13/#commen
t-1193243
http://www.gameinformer.com/b/features/archive/2016/04/28/understanding-autism-through-stardew-valley.aspx
https://web.archive.org/web/20170705220431/http://www.siliconera.com/2017/02/28/chucklefish-bringing-wargroove-
stardew-valley-pocket-rumble-switch/
https://www.destructoid.com/stardew-valley-is-headed-to-ios-this-month-and-then-android-526438.phtml
https://www.gamesradar.com/stardew-valley-update-16-drops-next-month-concernedape-confirms-now-back-to-the-grin
d/
https://www.polygon.com/2018/5/14/17352396/stardew-valley-ps-vita-release-date
https://www.pushsquare.com/games/ps4/stardew_valley
https://en.wikipedia.org/wiki/Stardew_Valley
https://www.pcgamer.com/stardew-valley-multiplayer-out-now-on-pc/
https://www.vg247.com/2017/08/01/stardew-valley-multiplayer-details-emerge-beta-test-coming-to-steam-at-the-end
-of-2017/
https://www.gamespot.com/articles/nintendo-switch-version-of-stardew-valley-releases/1100-6453715/
https://www.stardewvalleyconcert.com/
https://www.eurogamer.net/stardew-valleys-long-awaited-16-update-gets-march-release-date-on-pc
https://web.archive.org/web/20170718172936/http://www.gamasutra.com/view/news/267563/The_4_years_of_selfimposed
_crunch_that_went_into_Stardew_Valley.php
https://www.purexbox.com/games/xbox-one/stardew_valley
https://www.nintendolife.com/news/2023/10/stardew-valley-festival-of-seasons-concert-tour-announced
http://www.siliconera.com/2017/01/19/stardew-valley-vita-port-investigation/
http://www.pcgamesn.com/stardew-valley/stardew-valley-a-pc-harvest-moon-like-indie-game-rockets-to-the-top-of-s
team-top-sellers-chart
https://blog.playstation.com/2016/12/08/country-life-rpg-stardew-valley-coming-to-ps4-on-december-13/comment-pa
ge-2/#comments
http://www.gamasutra.com/view/news/331910/Stardew_Valley_dev_announces_decision_to_selfpublish.php
https://www.forbes.com/sites/mattperez/2017/01/03/30-under-30-in-games-2017-the-creators-and-founders-behind-ts
m-stardew-valley-and-cards-against-humanity/
https://www.king5.com/article/entertainment/television/programs/evening/eric-barone-stardew-valley/281-7a596cc6
-3c48-458b-bbda-672919d1c717
https://www.eurogamer.net/stardew-valley-creator-reveals-haunted-chocolatier?view=comments
https://www.bostonglobe.com/arts/2016/03/17/stardew-valley-you-reap-what-you-sow/OtevYl1D31M5x2KWGClHPK/story.h
tml
```

In [ ]:
```python
import re
def scrape_and_write_to_files(urls):
    file_id = 1
    for url in urls:
        try:
            # Fetch the HTML content of the URL
            response = requests.get(url)
            response.raise_for_status()  # Raise an exception for HTTP errors

        except Exception as e:
            print(e)

        # Parse the HTML of the link
        soup = BeautifulSoup(response.text, 'html.parser')

        # Check if the request was successful (status code 200)
        if response.status_code == 200:
            for script in soup(["script", "style"]):
                script.extract()
            # Extract text from the HTML
            # text = soup.get_text()

            # Extract text from all paragraphs on the webpage
            # Find all <p> elements without any classes specified
            paragraphs = soup.select('p:not([class])')

            # Initialize a variable to store the extracted text
            text = ''

            for paragraph in paragraphs:
                text += paragraph.get_text() + ' '

            # Creating the filename from the URL
            # Returns the website name except for the top level domain


            # Only if text exists, (if blank do nothing)
            if text:
                split_website = get_website_name(url).split('.')
                file_name = split_website[:-1]
                file_name = '-'.join(file_name)
                file_name = file_name.replace('.', '-')

                # Add the ID
                file_name += ('-raw-' + str(file_id))
                file_id    += 1

                # Add the .txt extention
                file_name += '.txt'

                # Add path
```

```
            file_path = "raw/" + file_name
            print(file_name)

            with open(file_path, 'w', encoding='utf-8') as file:
                file.write(text)

            print("Text scraped from %s and written to %s" % (url, file_name))

        else:
            print("Website fetch failed.")
```

```python
# Create "raw" folder on current directory:
import os

folder_name = "raw"

# Check if the folder exists
if not os.path.exists(folder_name):
    # If it doesn't exist, create it
    os.makedirs(folder_name)
    print("Folder %s created." %folder_name)
else:
    print("Folder %s exists." %folder_name)

scrape_and_write_to_files(urls)
```

```
Folder raw created.
eurogamer-raw-1.txt
Text scraped from https://www.eurogamer.net/stardew-valley-creator-reveals-haunted-chocolatier and written to e
urogamer-raw-1.txt
blog-us-playstation-raw-2.txt
Text scraped from https://blog.us.playstation.com/2016/12/08/country-life-rpg-stardew-valley-coming-to-ps4-on-d
ecember-13/#comment-1193243 and written to blog-us-playstation-raw-2.txt
gameinformer-raw-3.txt
Text scraped from http://www.gameinformer.com/b/features/archive/2016/04/28/understanding-autism-through-starde
w-valley.aspx and written to gameinformer-raw-3.txt
web-archive-raw-4.txt
Text scraped from https://web.archive.org/web/20170705220431/http://www.siliconera.com/2017/02/28/chucklefish-b
ringing-wargroove-stardew-valley-pocket-rumble-switch/ and written to web-archive-raw-4.txt
destructoid-raw-5.txt
Text scraped from https://www.destructoid.com/stardew-valley-is-headed-to-ios-this-month-and-then-android-52643
8.phtml and written to destructoid-raw-5.txt
gamesradar-raw-6.txt
Text scraped from https://www.gamesradar.com/stardew-valley-update-16-drops-next-month-concernedape-confirms-no
w-back-to-the-grind/ and written to gamesradar-raw-6.txt
polygon-raw-7.txt
Text scraped from https://www.polygon.com/2018/5/14/17352396/stardew-valley-ps-vita-release-date and written to
polygon-raw-7.txt
pushsquare-raw-8.txt
Text scraped from https://www.pushsquare.com/games/ps4/stardew_valley and written to pushsquare-raw-8.txt
en-wikipedia-raw-9.txt
Text scraped from https://en.wikipedia.org/wiki/Stardew_Valley and written to en-wikipedia-raw-9.txt
pcgamer-raw-10.txt
Text scraped from https://www.pcgamer.com/stardew-valley-multiplayer-out-now-on-pc/ and written to pcgamer-raw-
10.txt
vg247-raw-11.txt
Text scraped from https://www.vg247.com/2017/08/01/stardew-valley-multiplayer-details-emerge-beta-test-coming-t
o-steam-at-the-end-of-2017/ and written to vg247-raw-11.txt
gamespot-raw-12.txt
Text scraped from https://www.gamespot.com/articles/nintendo-switch-version-of-stardew-valley-releases/1100-645
3715/ and written to gamespot-raw-12.txt
eurogamer-raw-13.txt
Text scraped from https://www.eurogamer.net/stardew-valleys-long-awaited-16-update-gets-march-release-date-on-p
c and written to eurogamer-raw-13.txt
web-archive-raw-14.txt
Text scraped from https://web.archive.org/web/20170718172936/http://www.gamasutra.com/view/news/267563/The_4_ye
ars_of_selfimposed_crunch_that_went_into_Stardew_Valley.php and written to web-archive-raw-14.txt
purexbox-raw-15.txt
Text scraped from https://www.purexbox.com/games/xbox-one/stardew_valley and written to purexbox-raw-15.txt
nintendolife-raw-16.txt
Text scraped from https://www.nintendolife.com/news/2023/10/stardew-valley-festival-of-seasons-concert-tour-ann
ounced and written to nintendolife-raw-16.txt
siliconera-raw-17.txt
Text scraped from http://www.siliconera.com/2017/01/19/stardew-valley-vita-port-investigation/ and written to s
iliconera-raw-17.txt
pcgamesn-raw-18.txt
Text scraped from http://www.pcgamesn.com/stardew-valley/stardew-valley-a-pc-harvest-moon-like-indie-game-rocke
ts-to-the-top-of-steam-top-sellers-chart and written to pcgamesn-raw-18.txt
blog-playstation-raw-19.txt
Text scraped from https://blog.playstation.com/2016/12/08/country-life-rpg-stardew-valley-coming-to-ps4-on-dece
mber-13/comment-page-2/#comments and written to blog-playstation-raw-19.txt
forbes-raw-20.txt
Text scraped from https://www.forbes.com/sites/mattperez/2017/01/03/30-under-30-in-games-2017-the-creators-and-
founders-behind-tsm-stardew-valley-and-cards-against-humanity/ and written to forbes-raw-20.txt
king5-raw-21.txt
Text scraped from https://www.king5.com/article/entertainment/television/programs/evening/eric-barone-stardew-v
alley/281-7a596cc6-3c48-458b-bbda-672919d1c717 and written to king5-raw-21.txt
eurogamer-raw-22.txt
Text scraped from https://www.eurogamer.net/stardew-valley-creator-reveals-haunted-chocolatier?view=comments an
d written to eurogamer-raw-22.txt
```

In [ ]:
```python
import string

def ends_with_punctuation(text):
  # Get the last character of the text
  last_character = text[-1]

  # Check if  last character is a punctuation
  if last_character in string.punctuation:
    return True
  else:
    return False

# TEST
sentence1 = "^ This does end with a period."
sentence2 = "Nope, this one doesn't"
sentence3 = "Yup, this one doesn't, but ends with a question mark?"

print(ends_with_punctuation(sentence1))
print(ends_with_punctuation(sentence2))
print(ends_with_punctuation(sentence3))
```

```
True
False
True
```

```python
In [ ]:  def starts_with_char(text, char):
           text = text.strip()
           if text.startswith(char):
             return True
           else:
             return False

           # TEST
           sentence1 = "^ This does start with a ^."
           sentence2 = "Nope, this one doesn't"

           print(starts_with_char(sentence1, '^'))
           print(starts_with_char(sentence2, '^'))
```

```
True
False
```

```python
In [ ]:  def count_words(sentence):
           # Split sentence to word
           words = sentence.split()

           # Count number of words
           num_words = len(words)

           return num_words
```

```python
In [ ]:  def clean_up_file(text):

           # Remove ['num']
           filtered_sentence = re.sub(r'\[\d+\]', '', text)

           # Remove URLs
           filtered_sentence = re.sub(r'https?:\/\/\S*', '', filtered_sentence)

           # Remove whitespace
           filtered_sentence = re.sub(r'^\s*$', '', filtered_sentence)

           # Remove ©
           filtered_sentence = re.sub(r'©', '', filtered_sentence)

           # Sentence boundaries
           # filtered_sentence = re.sub(r'([.?!0-9])([A-Z])', r'\1 \2', filtered_sentence)

           # Remove leading whitespaces
           text_chunks = [chunk.strip() for chunk in filtered_sentence.splitlines()]

           cleaned_up_text = ""

           for i, chunk in enumerate(text_chunks):
             # Makes sure that doesn't start with '^' and chunk is larger than 5 words
             if not starts_with_char(chunk, '^') and count_words(chunk) > 5:
               # Adds a '.' if the sentence doesn't end in a punctuation
               if not ends_with_punctuation(chunk):
                 chunk += '.'
               # print(i+1, chunk)
               cleaned_up_text += chunk + ' '

           return cleaned_up_text
```

```python
In [ ]:  # Testing clean_up_file(text)

           test_file_folder = "raw/en-wikipedia-raw-9.txt"
           with open(test_file_folder, 'r') as file:
             file_contents = file.read()
             # print(file_contents)
             cleaned_up = clean_up_file(file_contents)

           print(cleaned_up)
```

Stardew Valley is a 2016 farm life sim game developed by Eric "ConcernedApe" Barone. Players take the role of a character who inherits their deceased grandfather's dilapidated farm in a place known as Stardew Valley. The game was released for Windows in February 2016 before being ported to other platforms. Stardew Valley is open-ended, allowing players to grow crops, raise livestock, fish, cook, mine, forage, and socialize with the townspeople, including the ability to marry and have children. It allows up to four players to play online together. Barone developed Stardew Valley by himself over four and a half years. He was heavily inspired by the Story of Seasons series, with additions to address some of the shortcomings of those games. He used it as an exercise to improve his own programming and game design skills. British studio Chucklefish approached Barone midway through development with the offer to publish the game, allowing him to focus more on completing it. Stardew Valley has been cited as one of the greatest video games of all time. Yasuhiro Wada, the creator of Story of Seasons, also praised the game for retaining the freedom that later entries of his series had lost. The game sold more than 30 million copies by 2024, making it one of the best-selling video games of all time. Stardew Valley is a farming simulation game primarily inspired by Story of Seasons, a series by Marvelous and previously known as Harvest Moon. At the start of the game, players create a character, who inherits a plot of land and a small house once owned by their grandfather in a small village called Pelican Town, located in the titular Stardew Valley. Players may select from several different farm types, each with a unique theme and different benefits and drawbacks. The farmland is initially overrun with boulders, trees, stumps, and weeds, and players must work to clear the debris in order to make way for crops, buildings, and various other placeable items. Players interact with non-player characters (NPCs) that inhabit the town, including engaging in relationships with them; this can culminat

e in marriage, which results in the NPC helping the player's character to tend the farm. After marriage, players may have the chance to conceive a child with their spouse, or adopt one in the case of same-sex marriage. Players can also engage in fishing, cooking, crafting, and exploring procedurally-generated caves with materials and ores to mine or creatures to combat. Players can take on various quests to earn additional money or complete specific collections of materials (called "bundles") to restore the town's Community Center or pay certain amounts of money to complete JojaMart bundles. Completing bundles rewards players with various items, including seeds and tools. Completing multiple bundles grants players access to previously restricted areas and game mechanics. All of these activities must be metered against the character's current health and exhaustion level, and the game's internal clock. The player can consume food that grants certain buffs that are useful in certain activities and situations. Food is also a source of replenishing health and energy, which allows the player to complete more tasks in a given day. The game uses a simplified calendar, each year having four 28-day months that represent each season, that determines which crops can be grown and which activities can be beneficial. Later in the game, players can restore a greenhouse and access a tropical island that can grow any crop regardless of the season. This island also contains a volcano and additional resources that can be utilized within the game. Stardew Valley, originally titled Sprout Valley, was created by American indie game designer Eric Barone, known profesionally as ConcernedApe. In 2011, Barone had graduated from the University of Washington Tacoma with a computer science degree but had not been able to get a job in the industry, instead working as an usher at the Paramount Theatre in Seattle. Looking to improve his computer skills for better job prospects, he came to the idea of crafting a game which would also pull in his artistic side. Growing up in the Pacific Northwest, Barone incorporated many elements of the region into the gameplay and art. Stardew Valley originally began as a modern fan-made alternative to the Story of Seasons series, as he felt the series had gotten "progressively worse" after Harvest Moon: Back to Nature. Unable to find a satisfactory replacement, Barone began to create a game similar to the series, stating that his intent was "to address the problems I had with [Story of Seasons]" and that "no title in the series ever brought it all together in a perfect way". Barone was also inspired by other games, including Animal Crossing, Rune Factory, Minecraft, and Terraria, adding features seen in those games such as crafting, quests, and combat. He was the sole developer on the game, creating all of the game's pixel art, music, sound effects, story and dialogue. Initially, Barone considered releasing Stardew Valley on Xbox Live Indie Games due to the ease of publishing on that platform, but found early on that his scope for it became much larger than originally anticipated. Barone publicly announced the game in September 2012, using Steam's Greenlight system to gauge interest on the game. After the game was shown a great deal of support from the community, Barone began working on it in full, engaging with Reddit and Twitter communities to discuss his progress and gain feedback on proposed additions. Shortly after the Greenlight period in 2013, he was approached by Finn Brice, director of Chucklefish, who offered to help publish the game on release. Chucklefish took over many of the non-development activities for Barone, such as site hosting and setting up his development wiki. Barone decided not to use Steam's early access feature for development, as he felt that it was not well suited to Stardew Valley. Barone spent four years working on the project, redoing it multiple times and frequently spending 10 hours or more a day working on it. He originally programmed it in C# using the Microsoft XNA framework, but later migrated to MonoGame in 2021, which, according to Barone, "futureproofs the game and allows mods to access more than 4 gigs of RAM". Barone aimed to give players the feeling of immersion in a small farming community, stating that he wanted Stardew Valley to be entertaining while also designing it to have "real-world messages". In contrast to earlier Story of Seasons games, which could end after two years of in-game time has passed, Barone kept Stardew Valley open-ended so that players would not feel rushed to try to complete everything possible. During development, Barone recognized that some players would attempt to figure out mechanically how to maximize their farm's yield and profit through spreadsheets and other tools, but hoped that most players would take the time to learn these on their own. To that end, he designed the cooking aspect of the game purposely not to be profitable, but instead to pay back in bonuses that aided exploration, farming, mining and fishing skills. Barone also opted not to include the butchering of farm animals for meat products, encouraging players to name and tend to each animal individually in staying with the feeling he wanted for the game. In April 2015, Barone announced he intended to release the game only once he felt it was feature complete, refusing to put the game onto the Early Access program, or accept any pre-sale payments. The game was released for Windows on February 26, 2016. Following its release, Barone continued to work on the game, taking feedback from the community and patching bugs, and stated plans to add in additional features at a later date. Barone anticipated adding in more end-game content, as well as ports for other platforms. Barone had stated that he initially planned a four-player cooperative mode to be released in the game at launch. In this mode, Barone planned that all players would share a common farm, enabling players to all do different tasks related to it, such as one player mining while others tend to different parts of the farm. The multiplayer feature supports both local area network and remote online connectivity. Barone had planned for public beta testing of the multiplayer feature in late 2017 for the Windows version, but was still working to improve the network code by early 2018. The multiplayer beta for Windows was released in April 2018, with it officially launching for all PC platforms on August 1, 2018. In December 2018, the multiplayer update was released for the Nintendo Switch. Mobile versions for iOS and Android were developed with help of The Secret Police, with the iOS version being released on October 24, 2018, and the Android version being released on March 14, 2019. Both versions include the ability for Windows, macOS, and Linux users to transfer progress to their device. In 2018, Barone stated his desire to assemble a team of developers to help continue further development of the game. By 2019, all versions of the game, except on mobile, were self-published by Barone. In December 2019, Stardew Valley was added to the Tesla Arcade, a Linux-based video game service incorporated within most models of Tesla electric cars. A cooperative board game adaptation, Stardew Valley: The Board Game, was released in February 2021. An official cookbook featuring adaptations of more than 50 in-game recipes was announced in August 2023 and is set to be released on May 14, 2024. ConcernedApe promoted its first Stardew Valley concert tour, Stardew Valley: Festival of Seasons, on October 10, 2023. In May 2016, Barone announced that publisher Chucklefish would help with non-English localizations, macOS, Linux, and console ports, and the technical aspects required for online co-operative play, allowing him to focus solely on the first major content update. The macOS and Linux ports were released on July 29, 2016. Ports for the PlayStation 4 and Xbox One were announced at E3 2016 in June. At the same event, Barone stated that a port for the Wii U was also to be released, although that version was later canceled in favor of a version for the Nintendo Switch. The PlayStation 4 and Xbox One versions were released respectively on December 13 and 14, 2016. The Switch version, ported by Sickhead Games, was released on October 5, 2017. In early 2017, Barone stated his intentions for the possibility of a PlayStation Vita port, which was later confirmed and released on May 22, 2018. Retail versions for the PlayStation 4 and Xbox One are published and distributed by 505 Games. A collector's edition released at the same time included a physical map of the game's world, a download code for the soundtrack, and a guide book. In December 2018, the publishing rights for all platforms outside of Android, iOS and Nintendo Switch reverted to Barone. This was followed by him regaining the rights for the Nintendo Switch release in October 2019, the official wiki in February 2021, the iOS release in December 2021, and the Android release in March 2022, ending Chucklefish's involvement. Stardew Valley has also seen an active modding community, with players adding various new features to the game. One such mod is Stardew Valley Expanded, which has over 2 million downloads. The majority of mods are only playable through the PC version of the game. Some updates to the game have included changes that make it easier to mod. Stardew Valley received "generally favorable" reviews, according to review aggregator Metacritic. It has been cited by several publications to be among the greatest video games of all time. Jesse Singal writing for The Boston Globe wrote that the game was "utterly compelling, lovingly crafted", and provided players with numerous varieties of activities to do without falling into a cycle of repetitive activities. Elise Favis of

Game Informer found that watching her autistic brother play Stardew Valley helped her understand his condition better, as the game provides enough structure of present events with enough of a view of future events to allow her brother to enjoy the game. Yasuhiro Wada, the creator of the Story of Seasons series that Stardew Valley was inspired by, stated that he was "very happy" with the game, as it has shown to him that Story of Seasons was not a forgotten series and continued in spirit. He also stated that the approach taken by Barone with Stardew Valley was able to retain the freedom that he had wanted to keep in the Story of Seasons series that had been lost in the later games, with more focus on animation and graphics. Gamasutra named Barone one of the top ten developers for 2016, identifying that he had "single-handedly" developed something that "breathed new life into a genre" otherwise dominated by the Story of Seasons series. In 2017, Forbes named Barone one of their "30 Under 30" people to watch in the area of video games by citing his commitment towards making Stardew Valley. Many media outlets praised Stardew Valley for its LGBT options and representation, notably including the option for the player character to marry one of 12 villagers regardless of gender. Gayming Magazine praised the option for a player to engage in same-sex marriage, while also criticizing the lack of racial diversity. Paste described the game as being anti-capitalist, citing the negative portrayal of the fictional Joja Corporation. Stardew Valley sold over 400,000 copies across Steam and GOG.com in two weeks, and more than a million within two months. Valve reported that Stardew Valley was in the top 24 revenue-generating games on Steam during 2016. Journalists noted that the gaming community had shown support for Barone for the game; while there had been some players who obtained the game illegally, these players were impressed with the game and stated they planned to purchase the game, while other players made offers to help pay for those who could not afford the game. An official sheet music book for piano and album was released in 2018 by label Materia Collective. By the end of 2017, Stardew Valley had sold more than 3.5 million copies across all platforms. The game was also the most downloaded on the Nintendo Switch for 2017, despite only being released in October of that year. Sensor Tower also estimated that for the game's first three weeks on the Apple App Store, it had earned more than US$1 million in revenue. By February 2024, Stardew Valley had sold over 30 million copies; 19 of which was on PC.

```python
# Writing x cleaned files
# ----------------------------

# Create new folder
proc_folder_name = 'process/'

# Check if the folder exists
if not os.path.exists(proc_folder_name):
  # If it doesn't exist, create it
  os.makedirs(proc_folder_name)
  print("Folder %s created." %proc_folder_name)
else:
  print("Folder %s exists." %proc_folder_name)

# Get the folder that holds the raw data
raw_dir = "raw/"

# Get all file names from raw_dir
file_names = os.listdir(raw_dir)

for name in file_names:
  curr_file = raw_dir + name
  # print(curr_file)

  # Opening each file in "raw" and reading its content
  with open(curr_file, 'r') as file:
    # print(curr_file)
    file_contents = file.read()
    clean_contents = clean_up_file(file_contents)
    # print(clean_contents)
    # print("-------------------------------------------------")

  # Writing the cleaned file to a different folder


  #Remove raw from name
  name = name.replace("-raw", "-clean")


  processed_file = proc_folder_name + name
  # print(processed_file)
  with open(processed_file, 'w', encoding='utf-8') as file:
    file.write(clean_contents)
    print("Text cleaned and written to %s" % (processed_file))
```

```
Folder process/ created.
Text cleaned and written to process/pcgamesn-clean-18.txt
Text cleaned and written to process/destructoid-clean-5.txt
Text cleaned and written to process/nintendolife-clean-16.txt
Text cleaned and written to process/eurogamer-clean-1.txt
Text cleaned and written to process/gamesradar-clean-6.txt
Text cleaned and written to process/polygon-clean-7.txt
Text cleaned and written to process/pushsquare-clean-8.txt
Text cleaned and written to process/blog-us-playstation-clean-2.txt
Text cleaned and written to process/gamespot-clean-12.txt
Text cleaned and written to process/eurogamer-clean-13.txt
Text cleaned and written to process/blog-playstation-clean-19.txt
Text cleaned and written to process/web-archive-clean-4.txt
Text cleaned and written to process/forbes-clean-20.txt
Text cleaned and written to process/eurogamer-clean-22.txt
Text cleaned and written to process/gameinformer-clean-3.txt
Text cleaned and written to process/purexbox-clean-15.txt
Text cleaned and written to process/en-wikipedia-clean-9.txt
Text cleaned and written to process/siliconera-clean-17.txt
Text cleaned and written to process/king5-clean-21.txt
Text cleaned and written to process/web-archive-clean-14.txt
Text cleaned and written to process/vg247-clean-11.txt
Text cleaned and written to process/pcgamer-clean-10.txt
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js