**LANDING PAGE**

<!DOCTYPE html>

<html>

<head>

  <title>Home Page</title>

</head>

<body data-gr-c-s-loaded="true">

          <header class="cd-header">

            <a  class="cd-menu-trigger" href="#main-nav">Menu<span></span></a>

          </header>

        </section> <!-- cd-intro -->

      </main>

      <nav id="main-nav">

        <ul id="menu\_display" style="list-style-type: none;">

          <li onclick="document.getElementById('login\_open').style.display='block'; document.getElementById('glitch\_icon').style.display='none'; $('#cd-main-content').removeClass('move-out');

          $('#main-nav').removeClass('is-visible');

          $('.cd-shadow-layer').removeClass('is-visible');"><a><span>Login</span></a></li>

          <li  onclick="$('#cd-main-content').removeClass('move-out');

          $('#main-nav').removeClass('is-visible');

          $('.cd-shadow-layer').removeClass('is-visible');"><a href="/upload\_file"><span>Upload/Download</span></a></li>

          <li onclick="$('#cd-main-content').removeClass('move-out');

          $('#main-nav').removeClass('is-visible');

          $('.cd-shadow-layer').removeClass('is-visible');"><a href="#About"><span>About</span></a></li>

          <li onclick="$('#cd-main-content').removeClass('move-out');

          $('#main-nav').removeClass('is-visible');

          $('.cd-shadow-layer').removeClass('is-visible');"><a href="#Contact"><span>Contact us</span></a></li>

      </ul>

        <a href="#0" class="cd-close-menu">Close<span></span></a>

      </nav>

  <div class="container1" id="login\_open" style="display: none;">

    <div class="frame">

      <div class="nav1">

        <ul class="links">

          <li1 class="signin-active"><a class="btn">Sign in</a></li1>

          <li1 class="signup-inactive" style="padding-left: 10px;"><a class="btn">Sign up </a></li1>

        </ul>

      </div>

      <div ng-app ng-init="checked = false">

                  <form class="form-signin" action="" method="post" name="form">

            <label for="username">Email </label>

            <input id="username\_login" class="form-styling" type="email" name="username" placeholder="Email"/>

            <label for="password">Password</label>

            <input id="password\_login" class="form-styling" type="password" name="password" placeholder="Password"/>

            <input type="checkbox" id="checkbox"/>

            <label for="checkbox" ><span class="ui"></span>Keep me signed in</label>

            <div class="btn-animate">

              <a  id ="signinbutton" class="btn-signin">Sign in</a>

            </div>

            <!-- <a id="btnLogout" class="btn-signin">Logout</a> -->

                  </form>

                  <form class="form-signup" action="" method="post" name="form">

            <label for="fullname">Full name</label>

            <input class="form-styling" type="text" name="fullname" placeholder="Full Name"/>

            <label for="email">Email</label>

            <input id="Email\_signup" class="form-styling" type="email" name="email" placeholder="E-mail"/>

            <label for="password">Password</label>

            <input id="Password\_signup" class="form-styling" type="password" name="password" placeholder="Password"/>

            <label for="confirmpassword">Confirm password</label>

            <input class="form-styling" type="password" name="confirmpassword" placeholder="Confirm Password"/>

            <a class="btn-signup" id="signup\_btn">Sign Up</a>

                  </form>

               </div>

        </div>

        <div class="forgot">

          <a href="Recover\_Password.html">Forgot your password?</a>

        </div>

    </div>

  </div>

  <div class="caption" id="glitch\_icon">

 <p class="glitch">B.A.G</p>

  </div>

</div>

<a name="About"></a>

<div style="color: #777;background-color:white;text-align:center;padding:50px 80px;text-align: justify;">

  <h3 style="text-align:center;">Introduction</h3>

  <p style="color: black;">This Project was started as a solution to the problem that data scientists spend majority of their time cleaning textual and numeric datasets before they are fit for use. We at B.A.G have tried to change this, create a tool for everyone to use for their various needs.</p>

</div>

<div class="bgimg-2" id="back-2">

  <div class="caption">

  <span class="border" style="background-color:transparent;font-size:25px;color: #f7f7f7;"></span>

  </div>

</div>

<div style="position:relative;">

  <div style="color:#ddd;background-color:#282E34;text-align:center;padding:50px 80px;text-align: justify;">

  <p>With this tool you'll be able to:</p>

  <p>-Generate an error report for the dataset</p>

  <p>-Which checks the spellings</p>

  <p>-Fill missing values with median imputation</p>

  <p>-Remove the duplicate row entries</p>

  <p>-Highlight the errors in the uploaded excel file</p>

</div>

</div>

<div style="position:relative;">

  <div style="color:#ddd;background-color:#282E34;text-align:center;padding:50px 80px;text-align: justify;">

  <p>Come give this tool and go, and find yourself making way more time for yourself!</p>

  </div>

</div>

<div class="bgimg-1" id="back-4">

  <div class="caption">

  <span class="border">Contact The Team</span>

  <a name="Contact"></a>

<div class="row" style="background-color:initial ;">

    <div class="column">

      <div class="card">

        <img src="../static/images/image(1).jpg" alt="Bhairavi" style="padding: 50px; height:400px; width:70%">

<script>

  // When the user clicks anywhere outside of the modal, close it

  window.onclick = function(event) {

      if (event.target == modal) {

        modal.style.display="none";

        glitch\_icon\_var.style.display="block";

      }

    }

setTimeout(function(){

document.getElementById('login\_open').style.display="block";

},10000);

  </script>

</body></html>

**UPLOAD/DOWNLOAD PAGE**

<div id="upload\_form" style="display: none; position: absolute; top: 300px; width: 98%; left: 10px;">

<form id="upload" action="/upload\_file" method="POST" enctype="multipart/form-data">

<!-- DragnDrop -->

<fieldset>

<legend style="color: #ffffff;"> Upload</legend>

<div>

<label style="color: #ffffff;" for="fileselect" >Files to upload:</label>

<input type="file" id="fileselect" name="excel\_file" multiple="false" onchange="ValidateSingleInput(this)" />

<div id="filedrag" >or drop files here</div>

</div>

<div id="submitbutton">

<button type="submit" >Upload Files</button>

</div>

</fieldset>

</form>

<div style="color: #ffffff;" id="messages">

<p>Status Messages</p>

</div>

<p style="display: block; font-family: Montserrat; color: white; text-align: center;"> To View the Error Report please use the link given below:</p>

<a id="Error\_" style="display: block; font-family: Montserrat; color: white; text-align: center;" href="/Error\_Report" > View Error Report</a>

</div>

<div id="Loader\_icon" class="dl">

<div class="dl\_\_container">

<div class="dl\_\_corner--top"></div>

<div class="dl\_\_corner--bottom"></div>

</div>

<div class="dl\_\_square"></div>

</div>

</body>

<style>

#filedrag.hover{

color:white;

border-color:white;

border-style: solid;

box-shadow: whitesmoke; }

</style>

// file selection

function FileSelectHandler(e) {

// cancel event and hover styling

FileDragHover(e)

// fetch FileList object

var files = e.target.files || e.dataTransfer.files;

// process all File objects

for (var i = 0, f; f = files[i]; i++) {

if(f.type=="application/vnd.openxmlformats-officedocument.spreadsheetml.sheet"|| f.type=="application/vnd.ms-excel")

{ParseFile(f)}

else{

alert("Sorry, " + f.name + " is invalid, allowed extensions are: " + \_validFileExtensions.join(", "));

} }}

<body>

<p style=" color:white; font-family:Montserrat;text-align:center;font-size: 40px;"> Error Report</p>

<p style=" color:white; font-family:Montserrat;text-align:center;font-size: 15px;"> This page contains a report on the Errors found in the uploaded file.</p>

<p style=" color:white; font-family:Montserrat;text-align:center;padding-top: 30px; font-size: 20px;"> Errors will be highlighted in the Downlaoded File according the Key given below:<

<!-- <pre style="font-size: large; position: absolute; top: 400px; left:600px ; color: white; "> -->

<pre style="font-size: large; text-align: center; color: white; ">

{{summary}}

</pre>

<p style=" color:white; font-family:Montserrat;text-align:center;font-size: 18px;">Postion of Each error present the the give Dataset</p>

<pre style="font-size: large; text-align: center; color: white; ">

<!-- <pre style="font-size: large; position: absolute; top: 600px; left:600px ; color: white; "> -->

{{text}}

</pre>

</div>

</body>

**Python**

def conversion(data):

data1=pd.concat([data[col].astype(str).str.lower() for col in data.columns], axis=1)

return data1

# converting words to num

# converting the int cols back to int from str

def wordtonum(data):

z=[]

for i in range (0,len(data.columns)):

c = 0

for j in range(0,len(data)):

if type(data.iat[j,i])==int:

c=c+1

else:

continue

if c>=(len(data)/2):

z.append(i)

# for k in range (0,len(data)):

# if type(data.iat[k,i]) == str :

# try:

# data.iat[k,i]=w2n.word\_to\_num(data.iat[k,i])

# except ValueError :

# pass

print(z)

return data,z

# converting words to num

# converting the int cols back to int from str

def wordtonum1(datat,z):

for i in z:

for k in range (0,len(datat))

if type(datat.iat[k,i]) == str :

try:

datat.iat[k,i]=w2n.word\_to\_num(datat.iat[k,i])

except ValueError :

pass

return datat

def spellcheck(datat):

set1=[]

workbook = load\_workbook(filename="C:/Users/Rakesh/Desktop/CS Project/App/Upload\_Folder/two.xlsx")

sheet = workbook.active

for i in range(0,len(datat)):

for j in range(0,len(datat.columns)):

if type(datat.iat[i,j]) == str :

if len(datat.iat[i,j].split()) == 1:

set1.append(datat.iat[i,j])

# elif len(sheet.cell\_value(i,j).split()) > 1:

words = open("C:/Users/Rakesh/Desktop/CS Project/App/static/Additional/spell\_words.txt").readlines()

words = [word.strip() for word in words]

d=[]

for i in range(0,len(set1)):

if set1[i] not in words:

d.append(set1[i])

for i in range(0,len(d)):

for j in range(0,len(datat)):

for k in range(0,len(datat.columns)):

if d[i]==datat.iat[j,k]:

sheet.cell(row=j+2, column=k+1).fill = PatternFill(fgColor='FFEE08', fill\_type = "solid")

workbook.save(filename="final.xlsx")

def missing\_summary(datat):

def missing\_values(datat):

workbook = load\_workbook(filename="C:/Users/Rakesh/Desktop/CS Project/App/final.xlsx")

sheet = workbook.active

missing\_values\_list = ["n/a", "na", "--","NaN", "nan","-"]

#print("\nMISSING VALUES IN:")

for i in range(0,len(missing\_values\_list)):

for j in range(0,len(datat)):

for k in range(0,len(datat.columns)):

if missing\_values\_list[i]==datat.iat[j,k]:

#print(j+2,"row and",k+1,"column")

string="Missing value in Row "+str(j+2)+" column "+str(k+1)+" \n"

f.write(string) # Writing Findings into text file

sheet.cell(row=j+2, column=k+1).fill = PatternFill(fgColor='FF0000', fill\_type = "solid")

try:

median = datat.iloc[:,k].median()

sheet.cell(row=j+2,column=k+1).value= median

datat.iat[j,k]=median

except(TypeError):

continue

workbook.save(filename="final.xlsx")

def run(data):

missing\_summary(data)

a,z=wordtonum(data)

b=conversion(a)

c=spaces(b)

d=wordtonum1(c,z)

e=duplicates\_removal(d)

e.to\_excel('C:/Users/Rakesh/Desktop/CS Project/App/Upload\_Folder/two.xlsx',index=False)

spellcheck(e)

missing\_values(e)

fix\_text()