Summary of Clinical Information Available

van de Vijver et al. (2002) Datasets
The van de Vijver et al. (2002) dataset had a number of clinical features available for each patient. This included tumor diameter (mm), ER status, grade 3 classes, age (years), lymph node (number positive,) and whether chemotherapy and/or hormonal therapy had been given. The breakdown of the above clinical features is broken down below.

<table>
<thead>
<tr>
<th>Treatment Given</th>
<th>Yes (1)</th>
<th>No (0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemotherapy</td>
<td>110</td>
<td>185</td>
</tr>
<tr>
<td>Hormonal Therapy</td>
<td>40</td>
<td>255</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ER status</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>(1) 226</td>
</tr>
<tr>
<td>Negative</td>
<td>(0) 69</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade 3 Classes</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poorly differentiated</td>
<td>(1) 119</td>
</tr>
<tr>
<td>Intermediate</td>
<td>(2) 101</td>
</tr>
<tr>
<td>Well differentiated</td>
<td>(3) 75</td>
</tr>
</tbody>
</table>

TCGA Colon Cancer (2012) Datasets
The TCGA Colon Cancer (2012) dataset contained the following clinical features for each of its patients: anatomic organ subdivision, tumor site, age at initial pathologic diagnosis, gender, histological type, history of colon polyps, lymphatic invasion present, number of first degree relatives with cancer diagnosis, tumor stage, and vascular invasion present. The breakdown of the spread of these clinical features is broken down below.

<table>
<thead>
<tr>
<th>Histological Type</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colon Mucinous Adenocarcinoma</td>
<td>(0) 24</td>
</tr>
<tr>
<td>Colon Adenocarcinoma</td>
<td>(1) 125</td>
</tr>
<tr>
<td>Not Available</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tumor Site</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>right colon</td>
<td>(1) 62</td>
</tr>
<tr>
<td>transverse colon</td>
<td>(2) 14</td>
</tr>
<tr>
<td>left colon</td>
<td>(3) 71</td>
</tr>
<tr>
<td>rectum</td>
<td>(4) 1</td>
</tr>
</tbody>
</table>

1 Note that the number in parentheses for each clinical feature corresponds to the numeric value present in the 2002_clinicalInfo7m.txt file.
2 Note that the number in parentheses for each clinical feature corresponds to the numeric value present in the ColonDf_clinicalInfo10m.txt file.
<table>
<thead>
<tr>
<th>Anatomic Organ Subdivision</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descending Colon</td>
<td>6</td>
</tr>
<tr>
<td>Hepatic Flexure</td>
<td>8</td>
</tr>
<tr>
<td>Rectosigmoid Junction</td>
<td>1</td>
</tr>
<tr>
<td>Ascending Colon</td>
<td>28</td>
</tr>
<tr>
<td>Splenic Flexure</td>
<td>2</td>
</tr>
<tr>
<td>Sigmoid Colon</td>
<td>63</td>
</tr>
<tr>
<td>Cecum</td>
<td>29</td>
</tr>
<tr>
<td>Transverse Colon</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Anatomic Organ Subdivision</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>28</td>
</tr>
<tr>
<td>Stage II</td>
<td>60</td>
</tr>
<tr>
<td>Stage III</td>
<td>39</td>
</tr>
<tr>
<td>Stage IV</td>
<td>23</td>
</tr>
<tr>
<td>Not Available</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Yes (1)</th>
<th>No (0)</th>
<th>Not Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>History of Colon Polyps</td>
<td>78</td>
<td>74</td>
</tr>
<tr>
<td>Lymphatic Invasion Present</td>
<td>77</td>
<td>70</td>
</tr>
<tr>
<td>Vascular Invasion Present</td>
<td>33</td>
<td>105</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of First Degree Relatives with Cancer Diagnosis</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>129</td>
</tr>
<tr>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (1)</td>
<td>77</td>
</tr>
<tr>
<td>Female (2)</td>
<td>75</td>
</tr>
</tbody>
</table>