

■ 図2-31：集計期間内での離脱顧客

```
In [32]: away_data = pd.merge(uriage_data, kokyaku_data, left_on="customer_name", right_on="顧客名", how="right")
         away_data[away_data["purchase_date"].isnull()]["顧客名", "メールアドレス", "登録日"]

Out[32]:
```

	顧客名	メールアドレス	登録日
2999	福井美樹	fukui_miki@example.com	2019-04-23

売上履歴と顧客台帳をライトジョインし、顧客台帳を主体として結合してみま
しょう。

集計期間内に購買を行っていない顧客は「購買日」等がNaNで結合されます。
それを条件にデータを抽出する事で、集計期間に購入を行っていない顧客を調べる事が可能です。



これで、第2章の10本ノックは終了です。

データ分析を行うにあたり、とても大事なデータ加工の基礎について学んできました。冒頭にもあります通り、データの状態を見極め、どのように加工するかを現場の方と密に相談しながらすめ、的確にデータ加工を行う必要があります。ここを疎かにすると後でとても痛い目にあうので、今回のような基礎的なクレンジングだけではなく、複雑なデータのクレンジングに挑戦してみてください。本章がそのきっかけとなれば幸いです。

次章からは、より高度で具体的な分析に踏み込んでいきます。

```
away_data = pd.merge(uriage_data, kokyaku_data, left_on="customer_name", right_on="顧客名", how="right")
away_data[away_data["purchase_date"].isnull()][["顧客名", "メールアドレス", "登録日"]]
```