Band gaps and material phases in dataset